


```

APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVIS WRIGHT TREMAINE LLP
STREET: 1501 Fourth Avenue, 2600 Century Square
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: PC compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/630,155
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Davison, Barry L.
REGISTRATION NUMBER: 47,309
REFERENCE/DOCKET NUMBER: 49321-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 628-7621
TELEFAX: 206 628-7699
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 79
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: HER-2 ECD antagonist
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-630-155-1

Query Match          94.4%; Score 355; DB 4; Length 7
Best Local Similarity 87.3%; Pred.No.3.3e-39;
Matches 69; Conservative 0; Mismatches 10; Indels

QY 1 GXHXXRPAAPVPEXRXQPAHPVLFLRPSWDVSAFYSLPLAP
DB 1 GTHSLRPAAPVPELRWQCPAHPVLFLRPSWDLVSAFYSLPLAP
QY 61 GRGXDPDAHVAVXLSRYEG 79
DB 61 GRGPDPAHVAVNLRYEG 79

RESULT 2
US-09-630-155-2
; Sequence 2, Application US/09630155
; Patent No. 6414130
; GENERAL INFORMATION:
; APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
; TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVIS WRIGHT TREMAINE LLP
; STREET: 1501 Fourth Avenue, 2600 Century Square
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: PC compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/630,155
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>

```

; Sequence 11, Application US/09147236A
; Patent No. 6316251
; GENERAL INFORMATION:
; APPLICANT: TONOUCHI, Naoto
; APPLICANT: TSUCHIDA, Takayasu
; APPLICANT: YOSHINAGA, Fumihiro
; APPLICANT: TAHARA, Naoki
; APPLICANT: HAYASHI, Takahisa
; TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE
; FILE REFERENCE: 6537-011-OPCT
; CURRENT APPLICATION NUMBER: US/09/147,236A
; CURRENT FILING DATE: 1999-04-08
; EARLIER APPLICATION NUMBER: PCT/JP97/03633
; EARLIER FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Acetobacter xylinum
; FEATURE:
; OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, c, or
; OTHER INFORMATION: t
US-09-147-236-11

Query Match 18.1%; Score 68; DB 4; Length 344;
Best Local Similarity 29.5%; Pred. No. 0.59;
Matches 18; Conservative 10; Mismatches 13; Indels 20; Gaps 3;

Qy 3 HXXPRPAVPVXXQXPAPVLSPL-----RPSWDXVSIFYSLPLSPTSXVI 55
Db 158 YAAAPQVATEVP--PQAPVAPVVAQAQVQRERPS-----LSPVTPPKPAV 204

Qy 56 S 56
205 S 205

Db

RESULT 5
US-09-522-474-11
; Sequence 11, Application US/09522474
; Patent No. 6573076
; GENERAL INFORMATION:
; APPLICANT: TONOUCHI, Naoto
; APPLICANT: TSUCHIDA, Takayasu
; APPLICANT: YOSHINAGA, Fumihiro
; APPLICANT: TAHARA, Naoki
; APPLICANT: HAYASHI, Takahisa
; TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE
; FILE REFERENCE: 6537-011-OPCT
; CURRENT APPLICATION NUMBER: US/09/522,474
; CURRENT FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US/09/147,236
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: PCT/JP97/03633
; PRIOR FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Acetobacter xylinum
; FEATURE:
; OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, c, or
; OTHER INFORMATION: t
US-09-522-474-11

Query Match 18.1%; Score 68; DB 4; Length 344;
Best Local Similarity 29.5%; Pred. No. 0.59;
Matches 18; Conservative 10; Mismatches 13; Indels 20; Gaps 3;

Qy 3 HXXPRPAVPVXXQXPAPVLSPL-----RPSWDXVSIFYSLPLSPTSXVI 55

Db 158 YAAAPQVATEVP--PQAPVAPVVAQAQVQRERPS-----LSPVTPPKPAV 204

Qy 56 S 56
205 S 205

Db

RESULT 6
US-09-252-991A-24956
; Sequence 24956, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24956
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24956

Query Match 17.8%; Score 67; DB 4; Length 158;
Best Local Similarity 35.7%; Pred. No. 0.31;
Matches 20; Conservative 8; Mismatches 26; Indels 2; Gaps 2;

Qy 7 PRP-AAVVPVXXQXPAPVLSPLRPSWDXVSIFYSLPL-APLSPTSXISPSV 60
Db 93 PPPFLSLPFLPLPLPLPLPLPLPLPLPLPLPLPLPLPLPLPLPLPL 148

RESULT 7
US-09-252-991A-29314
; Sequence 29314, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29314
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29314

Query Match 17.6%; Score 66; DB 4; Length 195;
Best Local Similarity 36.2%; Pred. No. 0.54;
Matches 25; Conservative 3; Mismatches 25; Indels 16; Gaps 4;

Qy 1 GXHXXPRPAVPVXXQXPAPVLSPLRPSWDXVSIFYSLPLSPTSXISPSV 59
Db 19 GLRHTAPRPAVP-----DQPAHFGPFAVRPRAARQAPGALPLR-----TPTL 64

Qy 60 VGR-GXDPD 67
Db 65 PGRHGRQPD 73

```
RESULT 8
US-08-430-286A-9
; Sequence 9, Application US/08430286A
; Patent No. 6225080
; GENERAL INFORMATION:
; APPLICANT: Uhl, George R.
; APPLICANT: Eppler, C. Mark
; APPLICANT: Wang, Jai-Bel
; TITLE OF INVENTION: Mu-Subtype Opioid Receptor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/430.286A
; FILING DATE: 28-APR-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Robinson, Joseph R.
; REGISTRATION NUMBER: 33,448
; REFERENCE/DOCKET NUMBER: 0646/1A843-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7700
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: OPB-R
US-08-430-286A-9
Query Match 17.0%; Score 64; DB 3; Length 440;
Best Local Similarity 37.8%; Pred. No. 2.7;
Matches 17; Conservative 4; Mismatches 14; Indels 10; Gaps 3;

QY 4 SXXPRPAAPVPPRX-----QXPAPHPVLSFLRPSWDXVSAFYSL 43
Db 26 TSSPAPTASPSAPSWTPSPRPGPAHP---FLQPPW--AVALWSL 65

RESULT 9
US-09-462-606-65
; Sequence 65, Application US/09462606
; Patent No. 6432408
; GENERAL INFORMATION:
; APPLICANT: MENG, XIANG-JIN
; APPLICANT: Emerson, Suzanne H.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: A SWINE HEPATITIS B VIRUS AND USES THEREOF
; FILE REFERENCE: 20264267US1
; CURRENT APPLICATION NUMBER: US/09/462.606
; CURRENT FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 60/053069
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: PC7/US98/14665
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Hepatitis E virus
US-09-462-606-65
Query Match 16.6%; Score 62.5; DB 4; Length 122;
Best Local Similarity 37.0%; Pred. NO. 0.86; Indels 1; Gaps 1;
Matches 17; Conservative 4; Mismatches 24; Indels 1; Gaps 1;

QY 13 PVEXRQXPAPHPVLSFLRPSWDXVSAFYSLPLAPLSPTSVXISPV 58
Db 65 PSPIFIQPTESLP-MSFHNFGLEFALDSRPAPLPLGVTSPSPAPPL 109

RESULT 10
US-09-252-991A-31949
; Sequence 31949, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31949
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31949
Query Match 16.5%; Score 62; DB 4; Length 495;
Best Local Similarity 35.9%; Pred. No. 5.6;
Matches 23; Conservative 9; Mismatches 26; Indels 6; Gaps 4;

QY 8 RPAAPVEXRQXPAPHPVLSFL- -RPSMDXVSAFYSLPLAPLSPTSVXISPVSVGRGXD 65
Db 284 RPAQPRPG-AGCPARPATGVARRPAGDALAA--SARPARI-PTLSRSPDAAPGPA 339
QY 66 PDAH 69
Db 340 PIPH 343

RESULT 11
US-08-611-107-4
; Sequence 4, Application US/08611107
; Patent No. 5801233
; GENERAL INFORMATION:
; APPLICANT: Haselkorn, Robert
; APPLICANT: Gornicki, Piotr
; TITLE OF INVENTION: NUCLEIC ACID COMPOSITIONS ENCODING
; TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE AND USES
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```


SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,107
FILING DATE: Concurrently Herewith
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US SN 07/956,700
FILING DATE: 02-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US93/09340
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US SN 08/422,560
FILING DATE: 14-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: ARCD:221
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 158 amino acids
TYPE: amino acid
STRANDEDNESS: linear
US-08-611-107-4

Query Match 16.2%; Score 61; DB 1; Length 158;
Best Local Similarity 33.9%; Pred. No. 1.9;
Matches 20; Conservative 6; Mismatches 29; Indels 4; Gaps 3;

QY 7 PRPAVPVXRQXPAPV--LSFLRPSWDXVSFAFVSLPLAPLSPTSVAISP-VSVGR 62
DB 56 PVPAPLPAPTAAAPPAGLGGKFLAITAPMVGTFYRAP-APREPPFVNVGDRIQVQ 113

RESULT 12
US-08-422-560A-4
Sequence 4, Application US/08422560A
Patent No. 5910626
GENERAL INFORMATION:
APPLICANT: Haselkorn, Robert
APPLICANT: Gornicki, Piotr
TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
METHODS FOR USE
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,560A
FILING DATE: 14-APR-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/956,700
FILING DATE: 02-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mark B.
REGISTRATION NUMBER: 37,259
REFERENCE/DOCKET NUMBER: ARCD:152/WIM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-418-3000

TELEFAX: 512-474-7577
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 158 amino acids
TYPE: amino acid
STRANDEDNESS: linear
US-08-422-560A-4

Query Match 16.2%; Score 61; DB 2; Length 158;
Best Local Similarity 33.9%; Pred. No. 1.9;
Matches 20; Conservative 6; Mismatches 29; Indels 4; Gaps 3;

QY 7 PRPAVPVXRQXPAPV--LSFLRPSWDXVSFAFVSLPLAPLSPTSVAISP-VSVGR 62
DB 56 PVPAPLPAPTAAAPPAGLGGKFLAITAPMVGTFYRAP-APREPPFVNVGDRIQVQ 113

RESULT 13
US-08-468-793-4
Sequence 4, Application US/08468793
Patent No. 6177267
GENERAL INFORMATION:
APPLICANT: Haselkorn, Robert
APPLICANT: Gornicki, Piotr
TITLE OF INVENTION: ACETYL-CoA CARBOXYLASE COMPOSITIONS AND
METHODS OF USE
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,793
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,560
FILING DATE: 14-APR-1995
APPLICATION NUMBER: US SN 07/956,700
FILING DATE: 02-OCT-1992
CLASSIFICATION: 800
APPLICATION NUMBER: PCT/US93/09340
FILING DATE: 30-SEP-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: ARCD:152/KIT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 158 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-793-4

Query Match 16.2%; Score 61; DB 3; Length 158;
Best Local Similarity 33.9%; Pred. No. 1.9;
Matches 20; Conservative 6; Mismatches 29; Indels 4; Gaps 3;

QY 7 BRPAAPVPRXQXPAPHPV--LSFLRPSWDXYSAFYSLPLAPLSPTSVXKISP-VSVGR 62
Db 56 PVPAPLPDAPTPRAAPAGPLGGKFLAITAPWGTETRAP-APPEPPFVNGDRIQVGQ 113

RESULT 14
US-09-252-991A-26580
Sequence 26580, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26580
LENGTH: 325
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26580

Query Match 16.2%; Score 61; DB 4; Length 325;
Best Local Similarity 29.9%; Pred. No. 4.5;
Matches 23; Conservative 6; Mismatches 30; Indels 18; Gaps 3;

QY 8 RPAAPVPRXQXPAPHPV-----XQYPAHPVL-----SFLRPSWDXYSAFYSLPLAPLSPT 51
Db 239 RPAAPVPRXQXPAPHPV-----XQYPAHPVL-----SFLRPSWDXYSAFYSLPLAPLSPT 51

QY 52 SVXISPVSVGRGKDPDA 68
Db 297 GALLSREAPSNGLSPPA 313

RESULT 15
US-08-615-170-21
Sequence 21, Application US/08615170
Patent No. 5776776
GENERAL INFORMATION:
APPLICANT: ORDAHL, Charles P.
APPLICANT: AZAKIE, Anthony
APPLICANT: MAR, Janet H.
APPLICANT: FARRANCE, Iain K.G.
APPLICANT: HALL, Deborah E.
APPLICANT: STEWART, Alexandre P.R.
APPLICANT: LARKIN, Sarah B.
TITLE OF INVENTION: DTEF-1 ISOFORMS AND USES THEREOF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/615,170
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01526

FILING DATE: 06-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/191,493
FILING DATE: 04-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 2307U-053120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-615-170-21

Query Match 16.2%; Score 61; DB 1; Length 432;
Best Local Similarity 32.8%; Pred. No. 6.4;
Matches 20; Conservative 9; Mismatches 22; Indels 10; Gaps 4;

QY 4 SXPRPAAPVPRXQXPAPHPVLSFLRPSWDXYSAFYSLPLAPLSPTSVXKISP 57
Db 155 SAAPRFWSQPIR--CQPGSQDIKFPAPVPIQPPMPBSLASYE-PLAPLPAAASAV-P 210

QY 59 V 58
Db 211 V 211

RESULT 16
US-08-615-170-19
Sequence 19, Application US/08615170
Patent No. 5776776
GENERAL INFORMATION:
APPLICANT: ORDAHL, Charles P.
APPLICANT: AZAKIE, Anthony
APPLICANT: MAR, Janet H.
APPLICANT: FARRANCE, Iain K.G.
APPLICANT: HALL, Deborah E.
APPLICANT: STEWART, Alexandre P.R.
APPLICANT: LARKIN, Sarah B.
TITLE OF INVENTION: DTEF-1 ISOFORMS AND USES THEREOF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/615,170
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01526
FILING DATE: 06-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/191,493
FILING DATE: 04-FEB-1994
CLASSIFICATION: 435

```
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Heslin, James M.
/ REGISTRATION NUMBER: 29,541
/ REFERENCE/DOCKET NUMBER: 2307U-053120
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 326-2400
/ TELEFAX: (415) 326-2422
/ INFORMATION FOR SEQ ID NO: 19:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 433 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-615-170-19

Query Match 16.2%; Score 61; DB 1; Length 433;
Best Local Similarity 32.8%; Pred. No. 6.5; Mismatches 22; Indels 10; Gaps 4;
Matches 20; Conservative 9;

Qy 4 SXXPRPAAVPVXRXQXP-AHPVLSFLRPSWDXV-----AFYSLPLAPLSETSVXI 57
Db 156 SAAPRFGSGFIP--QQPGSQDIKPPAQPAVPIQPPMPPLASYB-PLAPLPPAASAV-P 211

Qy 58 V 58
Db 212 V 212

RESULT 17
US-08-468-036-3
; Sequence 3, Application US/08468036
; Patent No. 5728806
; GENERAL INFORMATION:
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Hoekstra, Merl F.
; TITLE OF INVENTION: Materials and Methods Relating to Proteins that
; TITLE OF INVENTION: Interact with Casein Kinase I
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,036
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/184,605
; FILING DATE: 21-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5728806and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31784
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-036-3

Query Match 16.2%; Score 61; DB 1; Length 433;
Best Local Similarity 32.8%; Pred. No. 6.5; Mismatches 22; Indels 10; Gaps 4;
Matches 20; Conservative 9;

Qy 4 SXXPRPAAVPVXRXQXP-AHPVLSFLRPSWDXV-----AFYSLPLAPLSETSVXI 57
Db 156 SAAPRFGSGFIP--QQPGSQDIKPPAQPAVPIQPPMPPLASYB-PLAPLPPAASAV-P 211

Qy 58 V 58
Db 212 V 212

RESULT 18
US-08-376-843-3
; Sequence 3, Application US/08376843
; Patent No. 5846764
; GENERAL INFORMATION:
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Hoekstra, Merl F.
; TITLE OF INVENTION: Materials and Methods Relating to Proteins
; TITLE OF INVENTION: that Interact with Casein Kinase I
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/376,843
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/184,605
; FILING DATE: 21-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5846764and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31784
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-376-843-3

Query Match 16.2%; Score 61; DB 2; Length 595;
Best Local Similarity 34.7%; Pred. No. 9.6; Mismatches 35; Indels 8; Gaps 3;
Matches 26; Conservative 6;

Qy 3 HSXXPRPAAVPVXRXQXP-AHPVLSFLRPSWDXVSAFYSLPLAPLSETSVXI 61
Db 383 NSANPIPTVPFNFNAPPNMAFCVPMNHNLSGPAVSQPSFLPPAPL-PRDSGYSSSSPG 441

Qy 62 RGXD-----PDAHV 70
Db 442 QLLDILNSKKPDSNV 456

RESULT 19
US-08-376-843-3

Query Match 16.2%; Score 61; DB 2; Length 595;
Best Local Similarity 34.7%; Pred. No. 9.6; Mismatches 35; Indels 8; Gaps 3;
Matches 26; Conservative 6;

Qy 3 HSXXPRPAAVPVXRXQXP-AHPVLSFLRPSWDXVSAFYSLPLAPLSETSVXI 61
Db 383 NSANPIPTVPFNFNAPPNMAFCVPMNHNLSGPAVSQPSFLPPAPL-PRDSGYSSSSPG 441

Qy 62 RGXD-----PDAHV 70
Db 442 QLLDILNSKKPDSNV 456
```

```
US-09-023-905A-10
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18874
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18874
;
; Query Match 16.1%; Score 60.5; DB 4; Length 221;
; Best Local Similarity 28.8%; Pred. No. 3.3;
; Matches 23; Conservative 7; Mismatches 29; Indels 21; Gaps 5;
;
QY 3 HXSPRAAVPVRXQXPAHPVLSFLRPSWDXVSAFYSLPLAPLSPTSVXISPV 58
;
DB 32 HPSAAPAFAVPIARRRRPAP-RPA---LRPR-----PPGPAGP--VRLSPLPEPG 75
;
QY 59 -SVGRGXDPDAHVAVXLSRY 77
;
DB 76 PARGPGRQGRTRTAGAVHRH 95
;
RESULT 22
US-09-252-991A-26404
; Sequence 26404, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26404
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26404
;
; Query Match 16.1%; Score 60.5; DB 4; Length 221;
; Best Local Similarity 34.5%; Pred. No. 3.3;
; Matches 19; Conservative 4; Mismatches 15; Indels 17; Gaps 2;
;
QY 15 PXRXQXP-----AHPVLSFLRPSWDXVSAFYSLPLA-----PLSPTS 52
;
DB 25 PMLRPLPLTVLLRADRIEPLLPTRPTDALAAALTLPLASTILTSSPPLWPTS 79
;
RESULT 23
US-08-682-847-6
; Sequence 6, Application US/08682847
; Patent No. 5858989
; GENERAL INFORMATION:
; APPLICANT: BABIUK, LORNE
; APPLICANT: VAN DEN HURK, SYLVIA
; APPLICANT: ZAMS, TIM
; APPLICANT: FITZPATRICK, DAVID
; TITLE OF INVENTION: RECOMBINANT BOVINE HERPESVIRUS TYPE 1
; TITLE OF INVENTION: POLYPEPTIDES AND VACCINES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
```

```
US-09-023-905A-10
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18874
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18874
;
; Query Match 16.1%; Score 60.5; DB 4; Length 221;
; Best Local Similarity 28.8%; Pred. No. 3.3;
; Matches 23; Conservative 7; Mismatches 29; Indels 21; Gaps 5;
;
QY 3 HXSPRAAVPVRXQXPAHPVLSFLRPSWDXVSAFYSLPLAPLSPTSVXISPV 58
;
DB 32 HPSAAPAFAVPIARRRRPAP-RPA---LRPR-----PPGPAGP--VRLSPLPEPG 75
;
QY 59 -SVGRGXDPDAHVAVXLSRY 77
;
DB 76 PARGPGRQGRTRTAGAVHRH 95
;
RESULT 22
US-09-252-991A-26404
; Sequence 26404, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26404
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26404
;
; Query Match 16.1%; Score 60.5; DB 4; Length 221;
; Best Local Similarity 34.5%; Pred. No. 3.3;
; Matches 19; Conservative 4; Mismatches 15; Indels 17; Gaps 2;
;
QY 15 PXRXQXP-----AHPVLSFLRPSWDXVSAFYSLPLA-----PLSPTS 52
;
DB 25 PMLRPLPLTVLLRADRIEPLLPTRPTDALAAALTLPLASTILTSSPPLWPTS 79
;
RESULT 23
US-08-682-847-6
; Sequence 6, Application US/08682847
; Patent No. 5858989
; GENERAL INFORMATION:
; APPLICANT: BABIUK, LORNE
; APPLICANT: VAN DEN HURK, SYLVIA
; APPLICANT: ZAMS, TIM
; APPLICANT: FITZPATRICK, DAVID
; TITLE OF INVENTION: RECOMBINANT BOVINE HERPESVIRUS TYPE 1
; TITLE OF INVENTION: POLYPEPTIDES AND VACCINES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
```

```
US-09-023-905A-10
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18874
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18874
;
; Query Match 16.1%; Score 60.5; DB 4; Length 221;
; Best Local Similarity 28.8%; Pred. No. 3.3;
; Matches 23; Conservative 7; Mismatches 29; Indels 21; Gaps 5;
;
QY 3 HXSPRAAVPVRXQXPAHPVLSFLRPSWDXVSAFYSLPLAPLSPTSVXISPV 58
;
DB 32 HPSAAPAFAVPIARRRRPAP-RPA---LRPR-----PPGPAGP--VRLSPLPEPG 75
;
QY 59 -SVGRGXDPDAHVAVXLSRY 77
;
DB 76 PARGPGRQGRTRTAGAVHRH 95
;
RESULT 22
US-09-252-991A-26404
; Sequence 26404, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26404
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26404
;
; Query Match 16.1%; Score 60.5; DB 4; Length 221;
; Best Local Similarity 34.5%; Pred. No. 3.3;
; Matches 19; Conservative 4; Mismatches 15; Indels 17; Gaps 2;
;
QY 15 PXRXQXP-----AHPVLSFLRPSWDXVSAFYSLPLA-----PLSPTS 52
;
DB 25 PMLRPLPLTVLLRADRIEPLLPTRPTDALAAALTLPLASTILTSSPPLWPTS 79
;
RESULT 23
US-08-682-847-6
; Sequence 6, Application US/08682847
; Patent No. 5858989
; GENERAL INFORMATION:
; APPLICANT: BABIUK, LORNE
; APPLICANT: VAN DEN HURK, SYLVIA
; APPLICANT: ZAMS, TIM
; APPLICANT: FITZPATRICK, DAVID
; TITLE OF INVENTION: RECOMBINANT BOVINE HERPESVIRUS TYPE 1
; TITLE OF INVENTION: POLYPEPTIDES AND VACCINES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
```

```
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/682,847
; FILING DATE: 12-JUL-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PARK, FREDDIE K.
; REGISTRATION NUMBER: 35,636
; REFERENCE/DOCKET NUMBER: 29310-20005.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 417 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-682-847-6

Query Match 16.1%; Score 60.5; DB 2; Length 417;
Best Local Similarity 28.6%; Pred. No. 7.2;
Matches 18; Conservative 5; Mismatches 27; Indels 13; Gaps 2;

QY 1 GXHSXPRPAAPVPPRXQKPAHPVLSFLRPSWDXVSFAFYSLEPLSLPTSVMXISPVSV 60
DB 319 GGAEGEPKPG--PSPDADRP-----EGWPSLEATHPPPPATPATPAADPAVPVSV 365

QY 51 GRG 63
DB 366 GIG 368

RESULT 24
US-09-063-676-2
; Sequence 2, Application US/09063676
; Patent No. 6270800
; GENERAL INFORMATION:
; APPLICANT: Speaker, Tully J.
; APPLICANT: Clark, H. Fred
; APPLICANT: Moser, Charlotte A.
; APPLICANT: Offit, Paul A.
; APPLICANT: Campos, Manuel
; APPLICANT: Frenchick, Patrick J.
; TITLE OF INVENTION: AQUEOUS SOLVENT BASED ENCAPSULATION OF A
; TITLE OF INVENTION: BOVINE HERPES VIRUS TYPE-1 SUBUNIT VACCINE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pfizer Inc
; STREET: 235 East 42nd Street
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10017-5755
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/063,676
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 424
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Koller, Alan L. 37,371
; REGISTRATION NUMBER: PC9769A
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-573-2118
; TELEFAX: 212-573-1939
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 417 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-063-676-2

Query Match 16.1%; Score 60.5; DB 3; Length 417;
Best Local Similarity 28.6%; Pred. No. 7.2;
Matches 18; Conservative 5; Mismatches 27; Indels 13; Gaps 2;

QY 1 GXHSXPRPAAPVPPRXQKPAHPVLSFLRPSWDXVSFAFYSLEPLSLPTSVMXISPVSV 60
DB 319 GGAEGEPKPG--PSPDADRP-----EGWPSLEATHPPPPATPATPAADPAVPVSV 365

QY 51 GRG 63
DB 366 GIG 368

RESULT 25
US-09-252-991A-20121
; Sequence 20121, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20121
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (46)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
; US-09-252-991A-20121

Query Match 16.1%; Score 60.5; DB 4; Length 459;
Best Local Similarity 33.3%; Pred. No. 8.1;
Matches 24; Conservative 4; Mismatches 27; Indels 17; Gaps 4;

QY 1 GXHSXPRPAAPVPPRXQKPA-----HPVLSFLRPSWDXVSFAFYSLEPLSLPTSVMX 54
DB 231 GEHRGAARPARRVAPDRGQPATTPVEQPVRCARPRRGLHA--AQPVARTCAPT--- 285

QY 55 ISPVSVGRGXDP 66
DB 286 -----GRGERP 291

RESULT 26
US-09-535-008-63
; Sequence 63, Application US/09535008
; Patent No. 6465629
; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
```

```

; APPLICANT: Tavvagian, Sean V.
; APPLICANT: Teng, David H.-F.
; TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE
; TITLE OF INVENTION: AND OTHER CANCER TYPES
; FILE REFERENCE: 2318-259
; CURRENT APPLICATION NUMBER: US/09/535,008
; CURRENT FILING DATE: 2000-03-23
; EARLIER APPLICATION NUMBER: U.S. 60/125,806
; EARLIER FILING DATE: 1999-03-23
; SOFTWARE: PatentIn Ver. 2.0
; NUMBER OF SEQ ID NOS: 77
; SEQ ID NO 63
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-535-008-63

Query Match      16.1%; Score 60.5; DB 4; Length 543;
Best Local Similarity 27.9%; Pred. No. 9.9;
Matches 19; Conservative 5; Mismatches 19; Indels 25; Gaps 3;

QY 7 PRPAVP-----VPXRXQP--XPAHPVLSFLRPSWDXVSFAFYSLLPLAPLSPTSVXISPV 58
Db 313 PAPPVPPAASPVPMPPTQTSFGQQAQA-----PMPVPLHQKQSRITPI 355
QY 59 SVGRGXDP 66
Db 356 QKPRGLDP 363

RESULT 27
US-09-535-008-61
; Sequence 61, Application US/09535008
; Patent No. 6465629
; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavvagian, Sean V.
; APPLICANT: Teng, David H.-F.
; TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE
; TITLE OF INVENTION: AND OTHER CANCER TYPES
; FILE REFERENCE: 2318-259
; CURRENT APPLICATION NUMBER: US/09/535,008
; CURRENT FILING DATE: 2000-03-23
; EARLIER APPLICATION NUMBER: U.S. 60/125,806
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-535-008-61

Query Match      16.1%; Score 60.5; DB 4; Length 577;
Best Local Similarity 27.9%; Pred. No. 11;
Matches 19; Conservative 5; Mismatches 19; Indels 25; Gaps 3;

QY 7 PRPAVP-----VPXRXQP--XPAHPVLSFLRPSWDXVSFAFYSLLPLAPLSPTSVXISPV 58
Db 313 PAPPVPPAASPVPMPPTQTSFGQQAQA-----PMPVPLHQKQSRITPI 355
QY 59 SVGRGXDP 66
Db 356 QKPRGLDP 363

RESULT 28
US-09-077-940A-4
; Sequence 4, Application US/09077940A
; Patent No. 6576441
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru et al.
; TITLE OF INVENTION: NOVEL SENAPHORIN Z AND GENE ENCODING THE SAME

```

```

; FILE REFERENCE: 0020-4426P
; CURRENT APPLICATION NUMBER: US/09/077,940A
; CURRENT FILING DATE: 1998-06-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-077-940A-4

Query Match      16.1%; Score 60.5; DB 4; Length 888;
Best Local Similarity 36.2%; Pred. No. 18;
Matches 17; Conservative 5; Mismatches 18; Indels 7; Gaps 3;

QY 7 PRPAVPVPRXQXPAPHPVLSFLRP-SWDXVSFAFYSLLPLAPLSPTS 52
Db 708 PTPEQTPLPQKRLPTP-HPHPLALGPLAWD-----HGHPLLPASASS 748

RESULT 29
US-09-095-443-2
; Sequence 2, Application US/09095443
; Patent No. 6342593
; GENERAL INFORMATION:
; APPLICANT: Plowman, Gregory
; APPLICANT: Peles, Eior
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF ALP RELATED DISORDERS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,443
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/049,477
; FILING DATE: June 12, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 235/055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1274 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-095-443-2

Query Match      16.1%; Score 60.5; DB 4; Length 1274;
Best Local Similarity 29.4%; Pred. No. 28;
Matches 20; Conservative 8; Mismatches 27; Indels 13; Gaps 4;

QY 7 PRPA---AVPVPXRXQXPAPHPVLSF-LRPSWDXVSFAFYSLLPLAPLSPTSVXISPVSVGR 62

```

US 08 760 489-2

Db 664 PLPAHSGALPFPSPGPPPPHPPHPLAYGAPS-----TRPMGFOAAPLTIIRGPSSAGQ 715
Qy 63 GXDPDAHV 70
Db 716 ST-PSPHL 722

RESULT 30
US-08-760-489-2
; Sequence 2, Application US/08760489
; Patent No. 5830696
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: DIRECTED EVOLUTION OF THERMOPHILIC
; TITLE OF INVENTION: ENZYMES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: Fast-SEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,489
; FILING DATE: 05-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/008,311
; FILING DATE: 07-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/008001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1487 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-760-489-2

Query Match 16.1%; Score 60.5; DB 2; Length 1487;
Best Local Similarity 44.7%; Pred. No. 34;
Matches 21; Conservative 0; Mismatches 17; Indels 9; Gaps 3;
Qy 20 XPXA-HPVLSFLRPSWDXV-----SAFYSLP-LAPLSPTSXVISP 57
Db 1380 VPAPFHPFDLRLRSWSRVMSRKSSAQAKTAPLIPKSEISWPIPP 1426

Search completed: July 4, 2004, 04:09:23
Job time : 6.0763 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2004, 04:08:53 ; Search time 16.1807 Seconds
        (without alignments)
        1519.797 Million cell updates/sec

Title: US-09-506-079H-1
Perfect score: 376
Sequence: 1 GXHSXPRPAVPVEXRQP.....VGRGXDPDAHVAVXLSRYEG 79

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 31128316 residues
Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 90 summaries

Database : Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description
-----
1 356 94.7 79 12 US-10-344-470-1 Sequence 1, Appli
2 356 94.7 79 16 US-10-302-663-1 Sequence 1, Appli
3 352 93.6 419 12 US-10-344-470-2 Sequence 1, Appli
4 352 93.6 419 16 US-10-302-663-2 Sequence 2, Appli
5 78 20.7 616 9 US-09-925-300-1519 Sequence 1519, Ap
6 78 20.7 1579 16 US-10-309-290-82 Sequence 82, Appli
7 77 20.5 123 10 US-09-764-881-2743 Sequence 2743, Ap
8 76.5 20.3 555 16 US-10-437-963-203630 Sequence 203630,
9 75.5 20.1 427 16 US-10-437-963-203126 Sequence 203126,
10 74 19.7 1081 16 US-10-408-765A-1055 Sequence 1055, Ap
11 73 19.4 365 16 US-10-437-963-116009 Sequence 116009,
12 70 18.6 353 16 US-10-437-963-203523 Sequence 203523,
13 70 18.6 408 16 US-10-437-963-104664 Sequence 104664,
14 69.5 18.5 211 12 US-10-424-599-192240 Sequence 192240,
15 69.5 18.5 327 12 US-10-425-114-57214 Sequence 57214, A

```

```

16 69 18.4 141 16 US-10-437-963-184148 Sequence 184148,
17 69 18.4 2068 12 US-10-092-900A-220 Sequence 220, App
18 69 18.4 229 16 US-10-437-963-195562 Sequence 195562,
19 68.5 18.2 344 12 US-10-437-963-110471 Sequence 110471,
20 68.5 18.2 344 16 US-10-424-599-246855 Sequence 246855,
21 68 18.1 345 16 US-10-437-963-204246 Sequence 204246,
22 68 18.1 828 16 US-10-437-963-175075 Sequence 175075,
23 68 18.1 3503 13 US-10-108-605-237 Sequence 237, App
24 68 18.1 3503 14 US-10-174-677-2 Sequence 2, Appli
25 68 18.1 3503 14 US-10-437-963-176169 Sequence 3, Appli
26 67.5 18.0 407 14 US-10-080-170-588 Sequence 176189,
27 67.5 18.0 407 14 US-10-080-170-588 Sequence 588, App
28 67.5 18.0 407 16 US-10-080-170-588 Sequence 588, App
29 67 17.8 128 16 US-10-437-963-137244 Sequence 137244,
30 67 17.8 459 10 US-09-468-147-206 Sequence 206, App
31 67 17.8 459 10 US-09-468-147-206 Sequence 206, App
32 67 17.8 459 12 US-10-319-745-206 Sequence 207, App
33 67 17.8 459 12 US-10-319-745-207 Sequence 207, App
34 67 17.8 691 12 US-10-282-122A-67861 Sequence 67861, A
35 67 17.8 713 16 US-10-437-963-137248 Sequence 137248,
36 66.5 17.7 1127 16 US-10-437-963-124142 Sequence 124142,
37 66 17.6 271 16 US-10-437-963-167403 Sequence 167403,
38 65.5 17.4 108 16 US-10-437-963-111194 Sequence 111194,
39 65.5 17.4 147 16 US-10-437-963-185000 Sequence 185000,
40 65.5 17.4 581 16 US-10-437-963-127452 Sequence 127452,
41 65.5 17.4 917 12 US-10-412-699B-1657 Sequence 1657, App
42 65.5 17.4 917 15 US-10-374-780A-1624 Sequence 1624, Ap
43 65.5 17.4 917 16 US-10-437-963-141855 Sequence 141855,
44 65.5 17.4 1572 15 US-10-116-275-179 Sequence 179, App
45 65 17.3 359 12 US-10-425-114-42555 Sequence 42555, A
46 65 17.3 379 16 US-10-437-963-134191 Sequence 134191,
47 65 17.3 390 12 US-10-425-114-66070 Sequence 66070, A
48 65 17.3 564 16 US-10-437-963-130854 Sequence 130854,
49 64.5 17.2 107 12 US-10-424-599-173173 Sequence 173173,
50 64.5 17.2 146 16 US-10-437-963-106898 Sequence 106898,
51 64.5 17.2 166 16 US-10-437-963-202466 Sequence 202466,
52 64.5 17.2 448 16 US-10-437-963-199678 Sequence 199678,
53 64.5 17.2 554 16 US-10-437-963-174151 Sequence 174151,
54 64.5 17.2 1596 12 US-10-425-114-57832 Sequence 57832, A
55 64.5 17.2 2657 14 US-10-316-253-95 Sequence 95, Appli
56 64 17.0 154 16 US-10-437-963-134422 Sequence 134422,
57 64 17.0 193 16 US-10-437-963-113565 Sequence 113565,
58 64 17.0 215 14 US-10-156-761-12749 Sequence 12749, A
59 64 17.0 270 12 US-10-282-122A-66157 Sequence 66157, A
60 64 17.0 340 15 US-10-108-260A-3905 Sequence 3905, Ap
61 64 17.0 372 16 US-10-437-963-130371 Sequence 130371,
62 64 17.0 436 14 US-10-254-905-7 Sequence 7, Appli
63 64 17.0 440 10 US-09-966-782A-7 Sequence 7, Appli
64 64 17.0 440 14 US-10-225-567A-545 Sequence 545, App
65 64 17.0 824 16 US-10-437-963-142662 Sequence 142662,
66 64 17.0 1862 15 US-10-369-493-5974 Sequence 5974, Ap
67 64 17.0 97 16 US-10-437-963-201124 Sequence 201124,
68 63.5 16.9 132 10 US-09-468-147-203 Sequence 203, App
69 63.5 16.9 132 10 US-09-468-147-204 Sequence 204, App
70 63.5 16.9 132 12 US-10-319-745-203 Sequence 204, App
71 63.5 16.9 132 12 US-10-319-745-204 Sequence 204, App
72 63.5 16.9 132 16 US-10-437-963-203114 Sequence 203114,
73 63.5 16.9 223 16 US-10-437-963-127806 Sequence 127806,
74 63.5 16.9 480 15 US-10-310-154-658 Sequence 658, App
75 63.5 16.9 702 14 US-10-156-761-14859 Sequence 14859, A
76 63.5 16.9 108 16 US-10-437-963-200461 Sequence 200461,
77 63 16.8 135 12 US-10-424-599-248855 Sequence 248855,
78 63 16.8 316 16 US-10-437-963-116384 Sequence 116384,
79 63 16.8 316 16 US-10-437-963-147085 Sequence 147085,
80 63 16.8 322 16 US-10-437-963-127975 Sequence 127975,
81 63 16.8 356 16 US-10-437-963-194497 Sequence 194497,
82 63 16.8 366 16 US-10-437-963-118547 Sequence 118547,
83 63 16.8 374 12 US-10-425-114-64920 Sequence 64920, A
84 63 16.8 426 16 US-10-437-963-142229 Sequence 142229,
85 63 16.8 552 12 US-10-424-599-261717 Sequence 261717,
86 63 16.8 651 16 US-10-437-963-118908 Sequence 118908,
87 63 16.8 2448 12 US-10-310-172-48 Sequence 48, Appli
88 63 16.8 2448 12 US-10-310-172-48 Sequence 48, Appli

```



```
89 62.5 16.6 128 12 US-10-424-599-232671 Sequence 232671,
90 62.5 16.6 129 12 US-10-220-120-402 Sequence 402, App
ALIGNMENTS
RESULT 1
US-10-344-470-1
; Sequence 1, Application US/10344470
; Publication No. US20040052796A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; TITLE OF INVENTION: EXPRESSION OF HERSTATIN, AN ALTERNATIVE TO HER-2/NEU PRODUCT, IN C
; TITLE OF INVENTION: THAT EXPRESS EITHER P18SHER-2 OR THE EGF RECEPTOR INHIBITS RECE
; TITLE OF INVENTION: CELL GROWTH
; FILE REFERENCE: 49321-81
; CURRENT APPLICATION NUMBER: US/10/344,470
; CURRENT FILING DATE: 2003-06-09
; PRIOR APPLICATION NUMBER: PCT / US01/25502
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 09/638,834
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC FEATURE
; LOCATION: (2)..(21)
; OTHER INFORMATION: Applicants herein disclose Thr and Ser sequence variants at this
; OTHER INFORMATION: position
; NAME/KEY: MISC FEATURE
; LOCATION: (5)..(5)
; OTHER INFORMATION: Applicants herein disclose Leu and Pro sequence variants at this
; OTHER INFORMATION: position
; NAME/KEY: MISC FEATURE
; LOCATION: (6)..(6)
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
; OTHER INFORMATION: position
; NAME/KEY: MISC FEATURE
; LOCATION: (16)..(16)
; OTHER INFORMATION: Applicants herein disclose Leu and Gln sequence variants at this
; OTHER INFORMATION: position
; NAME/KEY: MISC FEATURE
; LOCATION: (18)..(18)
; OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at this
; OTHER INFORMATION: position
; NAME/KEY: MISC FEATURE
; LOCATION: (21)..(21)
; OTHER INFORMATION: Applicants herein disclose Gly, Asp, Ala and Val sequence variant
; OTHER INFORMATION: s at this position
; NAME/KEY: MISC FEATURE
; LOCATION: (36)..(36)
; OTHER INFORMATION: Applicants herein disclose Leu and Ile sequence variants at this
; OTHER INFORMATION: position
; NAME/KEY: MISC FEATURE
; LOCATION: (54)..(54)
; OTHER INFORMATION: Applicants herein disclose Pro and Arg sequence variants at this
; OTHER INFORMATION: position
; NAME/KEY: MISC FEATURE
; LOCATION: (64)..(64)
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
```

```
; OTHER INFORMATION: position
; NAME/KEY: MISC FEATURE
; LOCATION: (73)..(73)
; OTHER INFORMATION: Applicants herein disclose Asp and Asn sequence variants at thi
; OTHER INFORMATION: position
US-10-344-470-1
Query Match 94.7%; Score 356; DB 12; Length 79;
Best Local Similarity 100.0%; Pred. No. 1.6e-32;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GXHSXXPRAAVVPVXKXQXPAHPVLSFLRPSWDKXVSAPYSLPLAPLSPTSVXISPSV 60
DB 1 GXHSXXPRAAVVPVXKXQXPAHPVLSFLRPSWDKXVSAPYSLPLAPLSPTSVXISPSV 60
QY 61 GRXDPDAHVAVXLSRYEG 79
DB 61 GRXDPDAHVAVXLSRYEG 79
RESULT 2
US-10-302-663-1
; Sequence 1, Application US/10302663
; Publication No. US20040022785A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; TITLE OF INVENTION: EXPRESSION OF HERSTATIN, AN ALTERNATIVE HER-2/NEU PRODUCT, IN C
; TITLE OF INVENTION: EXPRESS EITHER P18SHER-2 OR THE EGF RECEPTOR INHIBITS RECEPTOR
; TITLE OF INVENTION: GROWTH
; FILE REFERENCE: 49321-73
; CURRENT APPLICATION NUMBER: US/10/302,663
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 09/638,834
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 1
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Homo Sapiens
; NAME/KEY: VARIANT
; LOCATION: 2
; OTHER INFORMATION: Applicants herein disclose Thr and Ser sequence variants at thi
; OTHER INFORMATION: position
; NAME/KEY: VARIANT
; LOCATION: 5
; OTHER INFORMATION: Applicants herein disclose Leu and Pro sequence variants at thi
; OTHER INFORMATION: position
; NAME/KEY: VARIANT
; LOCATION: 6
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at thi
; OTHER INFORMATION: position
; NAME/KEY: VARIANT
; LOCATION: 16
; OTHER INFORMATION: Applicants herein disclose Leu and Gln sequence variants at thi
; OTHER INFORMATION: position
; NAME/KEY: VARIANT
; LOCATION: 18
; OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at thi
; OTHER INFORMATION: position
; NAME/KEY: VARIANT
; LOCATION: 21
; OTHER INFORMATION: Applicants herein disclose Gly, Asp, Ala and Val sequence varia
; OTHER INFORMATION: position
; NAME/KEY: VARIANT
; LOCATION: 36
; OTHER INFORMATION: Applicants herein disclose Leu and Ile sequence variants at thi
; OTHER INFORMATION: position
; NAME/KEY: VARIANT
; LOCATION: 54
; OTHER INFORMATION: Applicants herein disclose Pro and Arg sequence variants at thi
```

```

1 FEATURE:
2 NAME/KEY: VARIANT
3 LOCATION: 64
4 OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
5 FEATURE:
6 NAME/KEY: VARIANT
7 LOCATION: 73
8 OTHER INFORMATION: Applicants herein disclose Asp and Asn sequence variants at this
9 US-10-302-663-1
10
11 Query Match 94.7%; Score 356; DB 16; Length 79;
12 Best Local Similarity 100.0%; Pred. No. 1.6e-32;
13 Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
14
15 QY 1 GXHSXXPRPAAVVPVXQXPXPAHPVLVSLRPSWDXVSFAFYSLPLAPLSPTSVXISPSV 60
16 DB 1 GXHSXXPRPAAVVPVXQXPXPAHPVLVSLRPSWDXVSFAFYSLPLAPLSPTSVXISPSV 60
17
18 QY 61 GRGXDPDAHVAVXLSRYEG 79
19 DB 61 GRGXDPDAHVAVXLSRYEG 79
20
21 RESULT 3
22 US-10-344-470-2
23 Sequence 2, Application US/10344470
24 Publication No. US20040052796A1
25 GENERAL INFORMATION:
26 APPLICANT: Clinton, Gail M.
27 TITLE OF INVENTION: EXPRESSION OF HERSTATIN, AN ALTERNATIVE TO HER-2/NEU PRODUCT, IN
28 TITLE OF INVENTION: THAT EXPRESS EITHER P185HER-2 OR THE EGF RECEPTOR INHIBITS RECEP
29 TITLE OF INVENTION: CELL GROWTH
30 FILE REFERENCE: 49321-81
31 CURRENT APPLICATION NUMBER: US/10/344,470
32 PRIOR FILING DATE: 2003-06-09 PCT / US01/25502
33 PRIOR APPLICATION NUMBER: PCT / US01/25502
34 PRIOR FILING DATE: 2001-08-14
35 PRIOR APPLICATION NUMBER: US 09/638,834
36 PRIOR FILING DATE: 2000-08-14
37 NUMBER OF SEQ ID NOS: 10
38 SOFTWARE: PatentIn version 3.1
39 SEQ ID NO 2
40 LENGTH: 419
41 TYPE: PRT
42 ORGANISM: Homo sapiens
43 FEATURE:
44 NAME/KEY: MISC FEATURE
45 LOCATION: (342)..(342)
46 OTHER INFORMATION: Applicants herein disclose Thr and Ser sequence variants at this
47 OTHER INFORMATION: position
48 FEATURE:
49 NAME/KEY: MISC FEATURE
50 LOCATION: (345)..(345)
51 OTHER INFORMATION: Applicants herein disclose Leu and Pro sequence variants at this
52 OTHER INFORMATION: position
53 FEATURE:
54 NAME/KEY: MISC FEATURE
55 LOCATION: (346)..(346)
56 OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
57 OTHER INFORMATION: position
58 FEATURE:
59 NAME/KEY: MISC FEATURE
60 LOCATION: (356)..(356)
61 OTHER INFORMATION: Applicants herein disclose Leu and Gln sequence variants at this
62 OTHER INFORMATION: position
63 FEATURE:
64 NAME/KEY: MISC FEATURE
65 LOCATION: (358)..(358)
66 OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at this
67 OTHER INFORMATION: position
68 FEATURE:
69 NAME/KEY: MISC FEATURE
70 LOCATION: (361)..(361)

```

NAME/KEY: VARIANT
LOCATION: 358
OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at this
FEATURE:
NAME/KEY: VARIANT
LOCATION: 361
OTHER INFORMATION: Applicants herein disclose Gly, Asp, Ala and Val sequence variants
OTHER INFORMATION: position
FEATURE:
NAME/KEY: VARIANT
LOCATION: 376
OTHER INFORMATION: Applicants herein disclose Leu and Ile sequence variants at this
FEATURE:
NAME/KEY: VARIANT
LOCATION: 394
OTHER INFORMATION: Applicants herein disclose Pro and Arg sequence variants at this
FEATURE:
NAME/KEY: VARIANT
LOCATION: 404
OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
FEATURE:
NAME/KEY: VARIANT
LOCATION: 413
OTHER INFORMATION: Applicants herein disclose Asp and Asn sequence variants at this
US-10-302-663-2

Query Match 93.6%; Score 352; DB 16; Length 419;
Best Local Similarity 98.7%; Pred. No. 2.7e-31;
Matches 78; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GXHXXPPAAVPPXQXPAPHPVLSPFLRPSWDXSAFYSLPLAPLSPTSVAISPSV 60
Db 341 GXHXXPPAAVPPXQXPAPHPVLSPFLRPSWDXSAFYSLPLAPLDPTSVAISPSV 400

Qy 61 GRGXDPAHVAVXLSRYEG 79
Db 401 GRGXDPAHVAVXLSRYEG 419

RESULT 5
US-09-925-300-1519
; Sequence 1519, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Steve Ruben,
; APPLICANT: Craig Rosen,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1519
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (262)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1519

Query Match 20.7%; Score 78; DB 9; Length 616;
Best Local Similarity 30.6%; Pred. No. 2.4;
Matches 22; Conservative 7; Mismatches 23; Indels 20; Gaps 3;

Qy 7 PRPAAVPPXQXPAPHPVLSPFLRPSWDXSAFYSLPLAPLSPTSVAISPSV 49
Db 323 PRPAVPPVSSAYPQGVHPAFLCAQYPSVTPPSLAATAVSFFVPSMAPIVHPYHTEPG 382

Qy 50 ---PTSXISPV 58
Db 383 LPLFTSVALLSSV 394

RESULT 6
US-10-309-290-82
; Sequence 82, Application US/10309290
; Publication No. US20040023241A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Anderson, David W.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Chillakuru, Rajeev A.
; APPLICANT: Edinger, Shomit R.
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Gorman, Linda
; APPLICANT: Gould-Rothberg, Bonnie E.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Jeffers, Michael E.
; APPLICANT: Ji, Weizhen
; APPLICANT: Li, Li
; APPLICANT: Maiyankar, Uriel M.
; APPLICANT: Miller, Charles E.
; APPLICANT: Murphey, Ryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Peyman, John A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Smithson, Glennda
; APPLICANT: Starling, Gary
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zhong, Haihong
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND MEET
; FILE REFERENCE: 21402-502A
; CURRENT APPLICATION NUMBER: US/10/309,290
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: 60/336,600
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,285
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/341,346
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/341,477
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/341,540
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/342,592
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/344,297
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/344,903
; PRIOR FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/373,288
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/380,981
; PRIOR FILING DATE: 2002-05-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 274
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 82
; LENGTH: 1579
; TYPE: PRI
; ORGANISM: Homo sapiens
US-10-309-290-82

```

Query Match      20.7%; Score 78; DB 16; Length 1579;
Best Local Similarity 30.6%; Pred. NC. 6.4;
Matches 25; Conservative 7; Mismatches 23; Indels 20; Gaps 3;

QY      7  PRPAAPVPRXRQPPXPAHPVL-----SFLRSDXVSAFYSLP-LAPUS----- 49
      ||||| ||||| ||||| ||||| :||| :||| :||| :||| :|||
1223  PRPAAPVPRSSAYPQGVHPAFLGAQVPSYVTPPSLAATAVFSFVPSNAPITVHFHYTEPG 1282

QY      50  ---PTSXVXISPV 58
      ||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db      1283  LPLPTSVALSIV 1294

```

RESULT 7

```

US-09-764-891-2743
/ Sequence 2743, Application US/09764891
/ Publication NO. US20030077808A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PC006
/ CURRENT APPLICATION NUMBER: US/09/764,891
/ CURRENT FILING DATE: 2001-01-17
/ Prior application data removed - consult PALM or file wrapper
/ NUMBER OF SEQ ID NOS: 10231
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 2743
/ LENGTH: 123
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (10)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (23)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ US-09-764-891-2743

```

```

Query Match      20.5%; Score 77; DB 10; Length 123;
Best Local Similarity 31.%; Pred. No. 0.55;
Matches 19; Conservative 11; Mismatches 22; Indels 8; Gaps 3;

QY 7 PRPAAVFVPRXQXPXPAHPVLSPRPSWDKVSIFYSLPLAPLSFTSVKISFVSVGRGXDP 66
47 PISLSLIP-HTHP-TSHP-----PSPEGSCFENRPLGLPPRHSHCPLAAGHPTP 98
Db

```

RESIST

```

US-10-437-963
; Sequence 203630, Application US/10437963
; Publication NO. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 203630
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:

```

```

; NAME/KEY: unsure
; LOCATION: (1)..{555}
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_98796C.1.pep
US-10-437-963-203630:

```

Query Match.

```

Query Match      20.1%; Score 76.5; DB 16; Length 555;
Best Local Similarity 36.9%; Pred. No. 3.1;
Matches 24; Conservative 5; Mismatches 31; Indels 5; Gaps 2;

QY      7  PRPAAVPVPRXKOPXPAHPVLSFLRSDWDXVSFAFYSGLPLAPLSFTSVKISPSVSVGRGXDP 66
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      120  PQAASG-QTTTPSPFSLLS---SPSPSGAELPLAPLSFTAPLCCFPPLGAAVPP 174

QY      67  DAHVA 71
          ||
Db      175  EXEVA 179

```

RESULT 9

```

US-10-437-963-203126
; Sequence 203126, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Bing
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 203126
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(427)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_9833C.1.psp
; US-10-437-963-203126

```

Query Match

```

Query Match      20.1%; Score 75.5; DB 16; Length 427;
Best Local Similarity 33.8%; Pred. NO. 3;
Matches 27; Conservative 10; Mismatches 30; Indels 13; Gaps 5;

QY      3 HXXPRPAAVPPVEXRXQXPAHPVLGSFLRPS---WDXVSAFY---SLPLAPLSPTSVXIS S6
DB      15 HQPWP-PAAPV-----AVPPSPVLGSL-PSAIKWDELTPFYRFNDIPSLPVVPSLPAA 67

QY      57 PVSVGRGXDPDAHVAVXLSR 76
DB      68 LLSLPRRRPPPAACAATVTR 87

RESULT 10
US-10-408-765A-1055
; Sequence 1055, Application US/10408765A
; Publication NO. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Robin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.

```

```
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408.765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1055
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1055

Query Match      19.7%; Score 74; DB 16; Length 1081;
Best Local Similarity 26.9%; Pred. No. 12;
Matches 25; Conservative 8; Mismatches 38; Indels 22; Gaps 4;

QY 7 PRPAAVPVXRXQXPAPHPVL-----SFLRPSWDXVSFYSILP-LAPLS----- 49
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 780 PRPAVPVPSAPYQGVHPAPLGAQYPSYVTPPSLAATAVSFFVESMAPITVHPYHTEPG 839
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 50 ---PTSVXISPVSVGRGXDDDAHVAVXLSRYEG 79
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 840 LPLPVSACE--LMGQGTSSVHPFPAIQG 870
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-10-437-963-116009
; Sequence 116009, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_19551C.1.pep
US-10-437-963-116009

Query Match      19.4%; Score 73; DB 16; Length 365;
Best Local Similarity 37.5%; Pred. No. 4.9;
Matches 18; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 3 HSXXPRPAAVPVXRXQXPAPHPVLSFLRPSWDXVSFYSILPAPLSP 50
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102 HFRVPRAPAPVPAPPAFTPIPTPPALAPPADVPPGMPPLVFTTP 149
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
US-10-437-963-203523
; Sequence 203523, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101977C.1.pep
US-10-437-963-203523

Query Match      18.6%; Score 70; DB 16; Length 408;
Best Local Similarity 33.3%; Pred. No. 12;
Matches 28; Conservative 5; Mismatches 27; Indels 24; Gaps 5;

QY 3 HSXXPRPA-----AVVPVXRXQ-----XPAPHPVLSFLRPSWDXVSFYSILPAPL 48
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 203523
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)....(353)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_9869C.1.pep
US-10-437-963-203523

Query Match      18.6%; Score 70; DB 16; Length 353;
Best Local Similarity 31.8%; Pred. No. 10;
Matches 21; Conservative 6; Mismatches 27; Indels 12; Gaps 2;

QY 7 PRPAAVPVXRXQXPAPHPVLSFLRPSWDXVSA-----FYSILPAPLSPTSVXISPVSV 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 83 PNPFPXPNPFPYHQSGAGLLPPPPPPAPPVSSPPLSTPHYSIPISPLSP-----PPPT 136
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 GRGXDP 66
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137 GRSP 142
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-10-437-963-104664
; Sequence 104664, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 104664
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)....(408)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101977C.1.pep
US-10-437-963-104664

Query Match      18.6%; Score 70; DB 16; Length 408;
Best Local Similarity 33.3%; Pred. No. 12;
Matches 28; Conservative 5; Mismatches 27; Indels 24; Gaps 5;

QY 3 HSXXPRPA-----AVVPVXRXQ-----XPAPHPVLSFLRPSWDXVSFYSILPAPL 48
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


APPLICANT: Catterton, Elina
APPLICANT: Leite, Mario W.
APPLICANT: Zhong, Haihong
APPLICANT: Alsobrook, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-290C
CURRENT APPLICATION NUMBER: US/10/092,900A
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: USN 60/274,322
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USN 60/283,675
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: USN 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: USN 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USN 60/274,191
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: USN 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: USN 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: USN 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: USN 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: USN 60/287,424
PRIOR FILING DATE: 2001-04-30
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 768
SEQ ID NO 220
LENGTH: 511
TYPE: PRT
ORGANISM: Homo sapiens
US-10-092-900A-220

Query Match 18.4%; Score 69; DB 12; Length 511;
Best Local Similarity 36.5%; Pred. No. 20;
Matches 19; Conservative 5; Mismatches 24; Indels 4; Gaps 1;

Qy 10 AAVPVXQXQPAHPVLSFLRPSWDXVSFAFVSLPLAPLSPVSVG 61
Db 215 AAAAAAGDPSIASFVSPCRP----VSSAARVPVPTSPSPASPSBITAG 262

RESULT 18
US-10-437-963-195562
Sequence 195562, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 195562
LENGTH: 2068
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure

LOCATION: (1)..(2068)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_91499C.1.pcp
US-10-437-963-195562

Query Match 18.4%; Score 69; DB 16; Length 2068;
Best Local Similarity 28.1%; Pred. No. 88;
Matches 25; Conservative 6; Mismatches 34; Indels 24; Gaps 4;

Qy 7 PRPAAVPVXRXQ-----XPAPHPVLSFLRPSWDXVSFAFVSLPLAPLSPVSVIS 56
Db 1551 PSPQAPATPPQYPATPPATPPATPPQAPLLAPSKSR-----ALPAPPATPATKKA 1605
Qy 57 PVSVGRGXDP-----DAHVAVXLSR 76
Db 1606 KVDAAXKNKDPGYDCTQBELDAHVASEVR 1634

RESULT 19

US-10-437-963-110471
Sequence 110471, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 110471
LENGTH: 229
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_14533C.1.pcp
US-10-437-963-110471

Query Match 18.2%; Score 68.5; DB 16; Length 229;
Best Local Similarity 33.3%; Pred. No. 9.5;
Matches 14; Conservative 7; Mismatches 20; Indels 1; Gaps 1;

Qy 7 PRPAAVPVXRXQXQXPAHPVLS-FLRPSWDXVSFAFVSLPLAP 47
Db 159 KPDPKPEPPKPKPEPEHPILDHFKKCKEDFFDHFHKKPVP 200

RESULT 20

US-10-424-599-246855
Sequence 246855, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 246855
LENGTH: 344
TYPE: PRT

```

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_64940C.1.pep
US-10-424-593-246855

Query Match      18.2%; Score 68.5; DB 12; Length 344;
Best Local Similarity 27.6%; Pred. No. 15;
Matches 16; Conservative 8; Mismatches 21; Indels 13; Gaps 1;

QY   13 PVPXRQXPDAH-----PVLSFLRPSWDXVSFAFVSLPLAPLSPTSVKISP 57
     |||:|||:|||||||:|||||:|||||:|||||:|||||:|||||:
Db    10 PLPLSLLPSPSHRSPSAKPFMAFTLTHTFSLPLSSPSSFHFPPLSLSPKPTSFNP 67

RESULT 21
US-10-437-963-204246
; Sequence 204246, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCES: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 204246
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: {1}..{345}
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_99352C.1.pep
US-10-437-963-204246

Query Match      18.1%; Score 68; DB 16; Length 345;
Best Local Similarity 28.8%; Pred. No. 17;
Matches 21; Conservative 10; Mismatches 26; Indels 16; Gaps 3;

QY   8 RPAAVPVRXXQPXAHPVLSELRP----SWDXVSFAFYSLPLAPLSPTSVKISPVSGR- 62
     |||:|||:|||||||:|||||:|||||:|||||:|||||:|||||:
Db    57 RPDSLETP---HPPPPPVLFLTKPRRSPSPAGHHRRPPPIINPSALPELPPTPVHRP 113

QY   63 -----GXDPD 67
Db    114 LPTPRLEWGRPPD 126

RESULT 22
US-10-437-963-175075
; Sequence 175075, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCES: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 175075
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: {1}..{345}
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_99352C.1.pep
US-10-437-963-204246

```


; CURRENT FILING DATE: 2002-06-19
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 3503
; TYPE: PRT
; ORGANISM: Drosophila
US-10-174-677-2

Query Match 18.1%; Score 68; DB 14; Length 3503;
Best Local Similarity 32.1%; Pred. No. 2e+02;
Matches 26; Conservative 10; Mismatches 29; Indels 16; Gaps 5;

QY 9 PAAVVPVXXQP-----XPAH-----PVSFLRPSWDXVAF-----YSLPLAPLSPTSVXISIP 57
DB 3411 PRAINLPMRLPPLHLSLAHAHLPRSPIGHEAGSFTSSAMSPFSPLSATRSPSISP 3470

QY 58 VSVGRGXDPDAHVA-VXLSRY 77
DB 3471 LGAG-----PPTLPHVSLPRH 3487

RESULT 25
US-10-174-677-3
; Sequence 3, Application US/10174677
; Publication No. US20030190704A1
; GENERAL INFORMATION:
; APPLICANT: Xie, Ting
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ANCHORING STEM CELLS IN A MICROENVIR
; FILE REFERENCE: 40716(IP-012)
; CURRENT APPLICATION NUMBER: US/10/174,677
; CURRENT FILING DATE: 2002-06-19
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 3503
; TYPE: PRT
; ORGANISM: Drosophila
US-10-174-677-3

Query Match 18.1%; Score 68; DB 14; Length 3503;
Best Local Similarity 32.1%; Pred. No. 2e+02;
Matches 26; Conservative 10; Mismatches 29; Indels 16; Gaps 5;

QY 9 PAAVVPVXXQP-----XPAH-----PVSFLRPSWDXVAF-----YSLPLAPLSPTSVXISIP 57
DB 3411 PRAINLPMRLPPLHLSLAHAHLPRSPIGHEAGSFTSSAMSPFSPLSATRSPSISP 3470

QY 58 VSVGRGXDPDAHVA-VXLSRY 77
DB 3471 LGAG-----PPTLPHVSLPRH 3487

RESULT 26
US-10-437-963-176169
; Sequence 176169, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 176169

Query Match 18.0%; Score 67.5; DB 16; Length 159;
Best Local Similarity 34.0%; Pred. No. 8.4;
Matches 18; Conservative 4; Mismatches 30; Indels 1; Gaps 1;

QY 7 PRPAVVPVXXQPXPAPHPVLSFLRPSWDXVSAFYSLPLAPLSPTSVXISIPVS 59
DB 50 PPPXPSPKRPSTPPPPPTFA-APPGWSTLPPRSPPPPLSPRSDVARAS 101

RESULT 27
US-10-080-170-588
; Sequence 588, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 588
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-588

Query Match 18.0%; Score 67.5; DB 14; Length 407;
Best Local Similarity 26.0%; Pred. No. 23;
Matches 25; Conservative 7; Mismatches 29; Indels 35; Gaps 3;

QY 8 RPAAVVPVXXQPXPAPHPVLSFLRPSWDXVSAFYSLPLAPLSPTSVXISIPVSGRXDPD 67
DB 41 RPAAVAVPT---PAPAREVFTSLKQLMTAAAPATRVFV-----VVGTVATGDGRQVD 90

QY 68 A-----HVAVXLSRYE 78
DB 91 GRDPATGESLSWYARDTDLGGVTWVHYHVAVYRYD 126

RESULT 28
US-10-080-170-588
; Sequence 588, Application US/10080170
; Publication No. US20040121322A9
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 588
; LENGTH: 407

; CURRENT FILING DATE: 2002-06-19
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 3503
; TYPE: PRT
; ORGANISM: Drosophila
US-10-174-677-2

Query Match 18.1%; Score 68; DB 14; Length 3503;
Best Local Similarity 32.1%; Pred. No. 2e+02;
Matches 26; Conservative 10; Mismatches 29; Indels 16; Gaps 5;

QY 9 PAAVVPVXXQP-----XPAH-----PVSFLRPSWDXVAF-----YSLPLAPLSPTSVXISIP 57
DB 3411 PRAINLPMRLPPLHLSLAHAHLPRSPIGHEAGSFTSSAMSPFSPLSATRSPSISP 3470

QY 58 VSVGRGXDPDAHVA-VXLSRY 77
DB 3471 LGAG-----PPTLPHVSLPRH 3487

RESULT 25
US-10-174-677-3
; Sequence 3, Application US/10174677
; Publication No. US20030190704A1
; GENERAL INFORMATION:
; APPLICANT: Xie, Ting
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ANCHORING STEM CELLS IN A MICROENVIR
; FILE REFERENCE: 40716(IP-012)
; CURRENT APPLICATION NUMBER: US/10/174,677
; CURRENT FILING DATE: 2002-06-19
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 3503
; TYPE: PRT
; ORGANISM: Drosophila
US-10-174-677-3

Query Match 18.1%; Score 68; DB 14; Length 3503;
Best Local Similarity 32.1%; Pred. No. 2e+02;
Matches 26; Conservative 10; Mismatches 29; Indels 16; Gaps 5;

QY 9 PAAVVPVXXQP-----XPAH-----PVSFLRPSWDXVAF-----YSLPLAPLSPTSVXISIP 57
DB 3411 PRAINLPMRLPPLHLSLAHAHLPRSPIGHEAGSFTSSAMSPFSPLSATRSPSISP 3470

QY 58 VSVGRGXDPDAHVA-VXLSRY 77
DB 3471 LGAG-----PPTLPHVSLPRH 3487

RESULT 26
US-10-437-963-176169
; Sequence 176169, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 176169

```
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-588

Query Match      18.0%; Score 67.5; DB 16; Length 407;
Best Local Similarity 26.0%; Pred. No. 23;
Matches 25; Conservative 7; Mismatches 29; Indels 35; Gaps 3;

QY 8 RPAAPVVRXQXPAPHPVLSFLRPSWDXVSAPYSLPLAPLSPTSVXISPVSVGKXDPD 67
DB 41 RPAAVAVPT---PAPAREVFTSLKQLWTAASPATRVFV-----VVGTVATGDRQVD 90
QY 68 A-----HVAVLSRYE 78
DB 91 GRDPAGESLWSYARDTDLGVTWVHYAVAVRYD 126

RESULT 29
US-10-437-963-137244
; Sequence 137244, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 137244
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_38745C.1.pap
US-10-437-963-137244
```

```
Query Match      17.8%; Score 67; DB 16; Length 128;
Best Local Similarity 35.7%; Pred. No. 7.5;
Matches 20; Conservative 6; Mismatches 20; Indels 10; Gaps 2;

QY 7 BPAAVPVVRXQXPAPHPVLSFLRPSWDXVSAPYSLPLAPLSPTSVXISPVSVGK 62
DB 17 PPPPVVFP-----PAISVTSVFP-----YSMTSSLPPSPRPPLPFPSPVIRR 62
```

```
RESULT 30
US-09-468-147-206
; Sequence 206, Application US/09468147A
; Publication No. US20030049601A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Eker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; FILE REFERENCE: 6232.US.P1
; CURRENT APPLICATION NUMBER: US/09/468,147A
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/173,141
; EARLIER FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: US 60/061,199
; EARLIER FILING DATE: 1997-10-15
```

```
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 206
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: CKSORF32M-3.pap
US-09-468-147-206

Query Match      17.8%; Score 67; DB 10; Length 459;
Best Local Similarity 31.3%; Pred. No. 30;
Matches 21; Conservative 7; Mismatches 29; Indels 10; Gaps 2;

QY 13 PVPXRXQXPAPHPVLSFLRPSWDXVSAPYSLPLAPLSPTS-----VXISPVSVGK 63
DB 67 PSPIFIQTPS-PPMSFHPNGLEALDSRPAPLAPLGVTSFSAAPLPVVDLPQLGLRRG 125
QY 64 XDPDAHV 70
DB 126 ADGTAEL 132

Search completed: July 4, 2004, 04:16:53
Job time : 18.1807 secs
```


APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVIS WRIGHT TREXANE LLP
STREET: 1501 Fourth Avenue, 2600 Century Square
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: PC compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Word

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/630,155
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Davison, Barry L.
REGISTRATION NUMBER: 47,309
REFERENCE/DOCKET NUMBER: 49321-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 628-7621
TELEFAX: 206 628-7699

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown

MOLECULE TYPE: polypeptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-630-155-2

Query Match 98.8%; Score 2220; DB 4; Length 419;
Best Local Similarity 97.4%; Pred. No. 1.4e-188;
Matches 408; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLLALLPPGAASCTGTDMKLRPLPASPETHLMDLRHLHYQGVVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGTDMKLRPLPASPETHLMDLRHLHYQGVVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVGYVLIHNOVQVPLQRLRVGTQLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVGYVLIHNOVQVPLQRLRVGTQLFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKQNQLA 180
DB 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKQNQLA 180
QY 181 LTLIDTNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVAGGCARCKGPLPTDCCHQOC 240
DB 181 LTLIDTNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVAGGCARCKGPLPTDCCHQOC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTIVCPHLNQEVTAEDGTQRCCKSKPCARXXXPRPAAPVPEVXKP 360
DB 301 YNYLSTDVGSCTIVCPHLNQEVTAEDGTQRCCKSKPCARXXXPRPAAPVPEVXKP 360
QY 361 XPAHPVLFLRPSWDVSAFYSLPLADPTSVXISPVSVGRGKDPDAHVAVKLSRYEG 419
DB 361 GPAHPVLFLRPSWDVSAFYSLPLADPTSVXISPVSVGRGKDPDAHVAVNLSRYEG 419

RESULT 2
US-09-146-283-4
Sequence 4, Application US/09146283
Patent No. 5976546

GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Ruegg, Curtis L.
APPLICANT: Wu, Hongyu
TITLE OF INVENTION: Immunostimulatory Compositions
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/146,283
FILING DATE: 03-SEPT-1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702
REFERENCE/DOCKET NUMBER: 7636-0010.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 782 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-09-146-283-4

Query Match 83.4%; Score 1873.5; DB 2; Length 782;
Best Local Similarity 89.3%; Pred. No. 1.6e-157;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;
QY 1 MELAALCRWGLLLALLPPGAASCTGTDMKLRPLPASPETHLMDLRHLHYQGVVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGTDMKLRPLPASPETHLMDLRHLHYQGVVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVGYVLIHNOVQVPLQRLRVGTQLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVGYVLIHNOVQVPLQRLRVGTQLFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKQNQLA 180
DB 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKQNQLA 180
QY 181 LTLIDTNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVAGGCARCKGPLPTDCCHQOC 240
DB 181 LTLIDTNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVAGGCARCKGPLPTDCCHQOC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTIVCPHLNQEVTAEDGTQRCCKSKPCAR---GXHXXPRPAAPVPEVX 357
DB 301 YNYLSTDVGSCTIVCPHLNQEVTAEDGTQRCCKSKPCAR---GXHXXPRPAAPVPEVX 357
QY 358 XQXPXA---HPVLSFLRPSWDVSAFYSLPLAP 387
DB 361 IQGFAGCKKIFGSLAPLPESFGDPSASNTAPLOP 394

RESULT 3
US-08-579-823A-4
; Sequence 4, Application US/08579823A
; Patent No. 6080409
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Rugg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Composition and Method
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,823A
; FILING DATE: 03-DEC-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 782 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-08-579-823A-4

Query Match 83.4%; Score 1873.5; DB 3; Length 782;
Best Local Similarity 89.3%; Pred. No. 1.6e-157;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;
QY 1 MELAALCRWGLLLALLPPGAASQVCTGTDKMLRLPASPEHLDMLRHLVQCGVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASQVCTGTDKMLRLPASPEHLDMLRHLVQCGVQGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGVYLLAHNQVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGVYLLAHNQVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQVQVPLQRLRIVRGTLQFEDNYALAVLDNG 180
DB 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQVQVPLQRLRIVRGTLQFEDNYALAVLDNG 180
QY 181 LTLIDNTRSRACHPCSPMCKSGRCSGESSDQSLTRTYCAGGCAKCKPLPTDCCHQEC 240
DB 181 LTLIDNTRSRACHPCSPMCKSGRCSGESSDQSLTRTYCAGGCAKCKPLPTDCCHQEC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMPNPGRYTTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMPNPGRYTTFGASCVTACP 300
QY 301 YNYLSTDVSGCTLVCPLEHQEVTAEDGTORCEKSKPCAR---GXHSXKPRPAAPVPR 357
DB 301 YNYLSTDVSGCTLVCPLEHQEVTAEDGTORCEKSKPCARVCVGLGMEHLREVRVTSAN 360

QY 358 XQXPA---HPVLSELRPSMDXVSFAFYSLPLAP 387
DB 361 IQEFAGCKKIFGSLAFPEFSDGDFASNTAPLQP 394
RESULT 4
US-09-344-195-4
; Sequence 4, Application US/09344195
; Patent No. 6210662
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Rugg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,195
; FILING DATE: 24-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 782 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-344-195-4

Query Match 83.4%; Score 1873.5; DB 3; Length 782;
Best Local Similarity 89.3%; Pred. No. 1.6e-157;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;
QY 1 MELAALCRWGLLLALLPPGAASQVCTGTDKMLRLPASPEHLDMLRHLVQCGVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASQVCTGTDKMLRLPASPEHLDMLRHLVQCGVQGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGVYLLAHNQVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGVYLLAHNQVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQVQVPLQRLRIVRGTLQFEDNYALAVLDNG 180
DB 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQVQVPLQRLRIVRGTLQFEDNYALAVLDNG 180
QY 181 LTLIDNTRSRACHPCSPMCKSGRCSGESSDQSLTRTYCAGGCAKCKPLPTDCCHQEC 240

```

Db 181 LTLIDNRSRACHPCSPMCKGSRGWCSESDCQSLTRTVCAAGCARCKGFLPTDCCHQC 240
Qy 241 AAGCTGPGCHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPGCHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 357
Db 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 357
Qy 358 XQXPXPA---HPVLSELRPSMDXVSFAFYSPLAP 387
Db 361 IQFAGCKKIFGSLAFLPESFDGDPASNTAPLQP 394

RESULT 5
US-08-467-083-68
; Sequence 68, Application US/08467083
; Patent No. 5726023
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,083
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US 08/414,417
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836 SEEDANBERRY
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-467-083-68

Query Match 83.4%; Score 1873.5; DB 1; Length 1255;
Best Local Similarity 89.3%; Pred. No. 2.9e-157;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

Qy 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLPASPETHLDMRLHLYGCGVVOGNL 60
Db 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLPASPETHLDMRLHLYGCGVVOGNL 60
Qy 61 ELTYLPTNASLSFLQDIEQVGYVLIHNOVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIEQVGYVLIHNOVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
Qy 121 DPLNNTTPTVGASPGSLRELQSLRTEILKGVLIQRPOLCYQDTILWKDIFHKNOLA 180

```

```

Db 121 DPLNNTTPTVGASPGSLRELQSLRTEILKGVLIQRPOLCYQDTILWKDIFHKNOLA 180
Qy 181 LTLIDNRSRACHPCSPMCKGSRGWCSESDCQSLTRTVCAAGCARCKGFLPTDCCHQC 240
Db 181 LTLIDNRSRACHPCSPMCKGSRGWCSESDCQSLTRTVCAAGCARCKGFLPTDCCHQC 240
Qy 241 AAGCTGPGCHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPGCHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 357
Db 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 357
Qy 358 XQXPXPA---HPVLSELRPSMDXVSFAFYSPLAP 387
Db 361 IQFAGCKKIFGSLAFLPESFDGDPASNTAPLQP 394

RESULT 6
US-08-414-417B-68
; Sequence 68, Application US/08414417B
; Patent No. 5801005
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,417B
; FILING DATE: 31-MAR-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-414-417B-68

Query Match 83.4%; Score 1873.5; DB 1; Length 1255;
Best Local Similarity 89.3%; Pred. No. 2.9e-157;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

Qy 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLPASPETHLDMRLHLYGCGVVOGNL 60
Db 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLPASPETHLDMRLHLYGCGVVOGNL 60
Qy 61 ELTYLPTNASLSFLQDIEQVGYVLIHNOVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIEQVGYVLIHNOVQVPLQRLIRVGTQLFEDNYALAVLDNG 120

```

QY 121 DPLNTPVTGASPGGLRLQLRSLSLTLKGGVLIQRPOLCYQDTILWKDIFHKNOLA 180
DB 121 DPLNTPVTGASPGGLRLQLRSLSLTLKGGVLIQRPOLCYQDTILWKDIFHKNOLA 180
QY 181 LTLIDNRSRACHPCSPCKGSRGWGSESDCQSLTRTVCAAGCARCKGPIPTDCCHQC 240
DB 181 LTLIDNRSRACHPCSPCKGSRGWGSESDCQSLTRTVCAAGCARCKGPIPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 357
DB 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 357
QY 358 XQXPXPA----HPVLSFLRPSWDXSAFYSPLAP 387
DB 361 IQEPAGCKKIFGSLAFPLPESFDGDPASNTAPLQP 394

RESULT 7

US-08-484-438-8
; Sequence 8, Application US/08484438
; Patent No. 5811098

GENERAL INFORMATION:
; APPLICANT: Ploewman, Gregory D.
; APPLICANT: Culouscou, Jean-Michel
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Siegall, Clay B.
; APPLICANT: Hellstr m, Ingegerd
; APPLICANT: Hellstr m, Karl E.

; TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,438
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/323,442
; FILING DATE: 14-OCT-1994
; APPLICATION NUMBER: US 08/150,704
; FILING DATE: 10-NOV-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/981,165
; FILING DATE: 24-NOV-1992
; CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-230
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid

STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-484-438-8

Query Match 83.4%; Score 1873.5; DB 2; Length 1255;
Best Local Similarity 89.3%; Pred. No. 2.9e-157;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASQTCTGTDMLKRLPASPTHLDMLRLHYGCGVQVQNL 60
DB 1 MELAALCRWGLLLALLPPGAASQTCTGTDMLKRLPASPTHLDMLRLHYGCGVQVQNL 60
QY 61 ELTVLPNASHLSPLQDIOEVQGVLIHAHQVQVLPQRLRIVRGTOIPEDNYALAVLDNG 120
DB 61 ELTVLPNASHLSPLQDIOEVQGVLIHAHQVQVLPQRLRIVRGTOIPEDNYALAVLDNG 120
QY 121 DPLNTPVTGASPGGLRLQLRSLSLTLKGGVLIQRPOLCYQDTILWKDIFHKNOLA 180
DB 121 DPLNTPVTGASPGGLRLQLRSLSLTLKGGVLIQRPOLCYQDTILWKDIFHKNOLA 180
QY 181 LTLIDNRSRACHPCSPCKGSRGWGSESDCQSLTRTVCAAGCARCKGPIPTDCCHQC 240
DB 181 LTLIDNRSRACHPCSPCKGSRGWGSESDCQSLTRTVCAAGCARCKGPIPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 357
DB 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 357
QY 358 XQXPXPA----HPVLSFLRPSWDXSAFYSPLAP 387
DB 361 IQEPAGCKKIFGSLAFPLPESFDGDPASNTAPLQP 394

RESULT 8

US-08-486-348A-68
; Sequence 68, Application US/08486348A
; Patent No. 5848538

GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.

; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,348A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:

TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-468-545B-68

Query Match 83.4%; Score 1873.5; DB 2; Length 1255;
Best Local Similarity 89.3%; Pred. No. 2.9e-157;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;
QY 1 MELAALCRWGILLALLPPGAASCTGCTDMKRLPASPETHLDMLRHLRYGQCQVQGNL 60
DB 1 MELAALCRWGILLALLPPGAASCTGCTDMKRLPASPETHLDMLRHLRYGQCQVQGNL 60
QY 61 ELTYLPTNASISFLQDIQEVQGVLIHNNVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
DB 61 ELTYLPTNASISFLQDIQEVQGVLIHNNVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPQCYQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPQCYQDTILWKDIFHKNQOLA 180
QY 181 LTLDITNRSRACHPCSPCKSRGWGESSDCOSLTRVTCAGGCARCKGPLPTDCCHEQC 240
DB 181 LTLDITNRSRACHPCSPCKSRGWGESSDCOSLTRVTCAGGCARCKGPLPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGYTFGASCVTACP 357
DB 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGYTFGASCVTACP 357
QY 358 QXPXPA---HPVLSFLRPSWDVXSAFYSLPLAP 387
DB 361 IQEFAGCKKIFGSLAFLEPESFDGDPASNTAPLOP 394

RESULT 11
US-08-356-786-2
Sequence 2, Application US/08356786
Patent No. 5877305
GENERAL INFORMATION:
APPLICANT: Huston, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. L.
APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08356786
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/831,967

FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-786-2

Query Match 83.4%; Score 1873.5; DB 2; Length 1255;
Best Local Similarity 89.3%; Pred. No. 2.9e-157;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;
QY 1 MELAALCRWGILLALLPPGAASCTGCTDMKRLPASPETHLDMLRHLRYGQCQVQGNL 60
DB 1 MELAALCRWGILLALLPPGAASCTGCTDMKRLPASPETHLDMLRHLRYGQCQVQGNL 60
QY 61 ELTYLPTNASISFLQDIQEVQGVLIHNNVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
DB 61 ELTYLPTNASISFLQDIQEVQGVLIHNNVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPQCYQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPQCYQDTILWKDIFHKNQOLA 180
QY 181 LTLDITNRSRACHPCSPCKSRGWGESSDCOSLTRVTCAGGCARCKGPLPTDCCHEQC 240
DB 181 LTLDITNRSRACHPCSPCKSRGWGESSDCOSLTRVTCAGGCARCKGPLPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGYTFGASCVTACP 357
DB 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGYTFGASCVTACP 357
QY 358 QXPXPA---HPVLSFLRPSWDVXSAFYSLPLAP 387
DB 361 IQEFAGCKKIFGSLAFLEPESFDGDPASNTAPLOP 394

RESULT 12
US-08-466-680B-68
Sequence 68, Application US/08466680B
Patent No. 6075122
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/466,680B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-466-680B-68

Query Match      83.4%; Score 1873.5; DB 3; Length 1255;
Best Local Similarity 89.3%; Pred. No. 2.9e-157;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

Qy 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPTHLDMLRLHYQGCVVQGNL 60
Db 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPTHLDMLRLHYQGCVVQGNL 60
Qy 61 ELTYLPTNASLSFLQDIOEVQGYVLIAHNQVQVPLQRLIRVGTOLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIOEVQGYVLIAHNQVQVPLQRLIRVGTOLFEDNYALAVLDNG 120
Qy 121 DPLNNTTPTVGTASPGGLRLQLRSLEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQLA 180
Db 121 DPLNNTTPTVGTASPGGLRLQLRSLEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQLA 180
Qy 181 LTLIDNRSRACHPCSPMCKGSRGWGSESDCSLRTVTCAGGCARCKGPLPTDCCHEQC 240
Db 181 LTLIDNRSRACHPCSPMCKGSRGWGSESDCSLRTVTCAGGCARCKGPLPTDCCHEQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAVFPVXR 357
Db 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAVFPVXR 357
Qy 358 QXPXPA---HPVLSFLRPSMDXVSFYSPLAP 387
Db 361 IQEFAGCKKIFGSLAFPEPESFDGDPASNTAPLQ 394

RESULT 14
US-09-811-115-3
; Sequence 3, Application US/09811115
; Patent No. 6632979
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-811-115-3

Query Match      83.4%; Score 1873.5; DB 4; Length 1255;
Best Local Similarity 89.3%; Pred. No. 2.9e-157;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

Qy 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPTHLDMLRLHYQGCVVQGNL 60
Db 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPTHLDMLRLHYQGCVVQGNL 60
Qy 61 ELTYLPTNASLSFLQDIOEVQGYVLIAHNQVQVPLQRLIRVGTOLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIOEVQGYVLIAHNQVQVPLQRLIRVGTOLFEDNYALAVLDNG 120
Qy 121 DPLNNTTPTVGTASPGGLRLQLRSLEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQLA 180
Db 121 DPLNNTTPTVGTASPGGLRLQLRSLEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQLA 180
Qy 181 LTLIDNRSRACHPCSPMCKGSRGWGSESDCSLRTVTCAGGCARCKGPLPTDCCHEQC 240
Db 181 LTLIDNRSRACHPCSPMCKGSRGWGSESDCSLRTVTCAGGCARCKGPLPTDCCHEQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

RESULT 13
US-09-527-487-2
; Sequence 2, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-527-487-2

Query Match      83.4%; Score 1873.5; DB 4; Length 1255;
Best Local Similarity 89.3%; Pred. No. 2.9e-157;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

Qy 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPTHLDMLRLHYQGCVVQGNL 60

```

QY 301 YNYLSTDVGSCTLVCPHNOEVTAEDGTQRCCKSKPCAR----GXHSXXPRPAVVPVEXR 357
Db 301 YNYLSTDVGSCTLVCPHNOEVTAEDGTQRCCKSKPCARVCGYGLGMEHLREVRVTSAN 360
QY 358 XQXPXA---HPVLSFLRPSNDXVSFAFYSPLAP 387
Db 361 IOEPAGCKKIFGSLAFPLPESFDGDPASNTAPLOP 394

RESULT 15

US-09-354-533-68
; Sequence 68, Application US/09354533
; Patent No. 6664370
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; DISCLOSURE: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; City: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/354,533
; FILING DATE: 15-Jul-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 68:

Query Match 83.4%; Score 1873.5; DB 4; Length 1255;
Best Local Similarity 89.3%; Pred. No. 2.9e-157;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;
QY 1 MELALCFWGLLALLPPGASTVCTGTMKRLPASPETHLDMRLHYGCGVQGNL 60
Db 1 MELALCFWGLLALLPPGASTVCTGTMKRLPASPETHLDMRLHYGCGVQGNL 60
QY 61 ELTYLPTNASLFLQDIEVQGYVLLAHNQVRQVPLQRLRIVRGTLFEDNYALAVLNG 120
Db 61 ELTYLPTNASLFLQDIEVQGYVLLAHNQVRQVPLQRLRIVRGTLFEDNYALAVLNG 120
QY 121 DPLNNTTPTGASPGGLREHLSRLTELLKGGVLIQRNPQCYQDTILWKDIFHKNOLA 180
Db 121 DPLNNTTPTGASPGGLREHLSRLTELLKGGVLIQRNPQCYQDTILWKDIFHKNOLA 180
QY 181 LTLIDTNRSRACHPCSPMKGSRCKWGESSEDCQSLTRTVACGGCARCKGPLPTDCCHEQC 240
Db 181 LTLIDTNRSRACHPCSPMKGSRCKWGESSEDCQSLTRTVACGGCARCKGPLPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLAHFNHSGICELHCPALVTNTDTFESMPNPEGRYTFGASCVTACP 300

Db 241 AAGCTGPKHSDCLAHFNHSGICELHCPALVTNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNOEVTAEDGTQRCCKSKPCAR----GXHSXXPRPAVVPVEXR 357
Db 301 YNYLSTDVGSCTLVCPHNOEVTAEDGTQRCCKSKPCARVCGYGLGMEHLREVRVTSAN 360
QY 358 XQXPXA---HPVLSFLRPSNDXVSFAFYSPLAP 387
Db 361 IOEPAGCKKIFGSLAFPLPESFDGDPASNTAPLOP 394

RESULT 16

US-08-422-108-1
; Sequence 1, Application US/08422108
; Patent No. 6015567
; GENERAL INFORMATION:
; APPLICANT: Hudziak, Robert M.
; APPLICANT: Shepard, H. Michael
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,108
; FILING DATE: 14-Apr-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/355460
; FILING DATE: 13-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/048346
; FILING DATE: 15-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/354319
; FILING DATE: 19-MAY-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 554C2D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 624 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-422-108-1

Query Match 78.5%; Score 1764.5; DB 3; Length 624;
Best Local Similarity 88.7%; Pred. No. 5.4e-148;
Matches 331; Conservative 3; Mismatches 32; Indels 7; Gaps 2;
QY 22 STQVCTGTMKRLPASPETHLDMRLHYGCGVQGNLELTYPNNAISLFLQDIEVQ 81
Db 1 STQVCTGTMKRLPASPETHLDMRLHYGCGVQGNLELTYPNNAISLFLQDIEVQ 60
QY 82 GYVLIHNOVRQVPLQRLRIVRGTLFEDNYALAVLNGDPLNNTTPTGASPGGLRELQ 141
Db 61 GYVLIHNOVRQVPLQRLRIVRGTLFEDNYALAVLNGDPLNNTTPTGASPGGLRELQ 120

142 LRSLSLTKGVLQIQRNPOLCYQDTILWKDIFHKNQALTLIDTNRSPACHPCSPMCKG 201
121 LRSLSLTKGVLQIQRNPOLCYQDTILWKDIFHKNQALTLIDTNRSPACHPCSPMCKG 180
202 SRCWGESSEDCOSLTRVTCAGGACRCKGPLEPTDCCHQCAAGCTGPKHSDCLACLHFNHS 261
181 SRCWGESSEDCOSLTRVTCAGGACRCKGPLEPTDCCHQCAAGCTGPKHSDCLACLHFNHS 240
262 GICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLTSTDVGSCTLVCPPLHNOE 321
241 GICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLTSTDVGSCTLVCPPLHNOE 300
322 VTAEQGTQRCCKSKPCAR---GXHSXXPRPAAVVPVXRXQXPAA---HPVLSFLRPSW 374
301 VTAEQGTQRCCKSKPCARVYGLGMEHLREVRAVTSANIQEFAGCKKIFGSLAFLEPSP 360
375 DXVSIFYSLPLAP 387
361 DGDPASNTAPLOP 373

RESULT 17

US-08-422-734-1
; Sequence 1, Application US/08422734
; Patent No. 6333159
; GENERAL INFORMATION:
; APPLICANT: Hudziak, Robert M.
; APPLICANT: Shepard, H. Michael
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,734
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/422108
; FILING DATE: 14-Apr-1995
; APPLICATION NUMBER: 08/355460
; FILING DATE: 13-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/048346
; FILING DATE: 15-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/354319
; FILING DATE: 19-MAY-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 554C2D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 624 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear

US-08-422-734-1

Query Match 78.5%; Score 1764.5; DB 4; Length 624;
Best Local Similarity 88.7%; Pred. No. 5.4e-148;
Matches 331; Conservative 3; Mismatches 32; Indels 7; Gaps 2;
QY 22 STQVCTGTDMKRLRPASPETHLDMLRHLYQGCVVQGNLELTYLPTNASLSFLQDIOEVQ 81
DB 1 STQVCTGTDMKRLRPASPETHLDMLRHLYQGCVVQGNLELTYLPTNASLSFLQDIOEVQ 60
QY 82 GYVLIAHNOVRQVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTPVTGASPGGLRELO 141
DB 61 GYVLIAHNOVRQVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTPVTGASPGGLRELO 120
QY 142 LRSLSLTKGVLQIQRNPOLCYQDTILWKDIFHKNQALTLIDTNRSPACHPCSPMCKG 201
DB 121 LRSLSLTKGVLQIQRNPOLCYQDTILWKDIFHKNQALTLIDTNRSPACHPCSPMCKG 180
QY 202 SRCWGESSEDCOSLTRVTCAGGACRCKGPLEPTDCCHQCAAGCTGPKHSDCLACLHFNHS 261
DB 181 SRCWGESSEDCOSLTRVTCAGGACRCKGPLEPTDCCHQCAAGCTGPKHSDCLACLHFNHS 240
QY 262 GICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLTSTDVGSCTLVCPPLHNOE 321
DB 241 GICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLTSTDVGSCTLVCPPLHNOE 300
QY 322 VTAEQGTQRCCKSKPCAR---GXHSXXPRPAAVVPVXRXQXPAA---HPVLSFLRPSW 374
DB 301 VTAEQGTQRCCKSKPCARVYGLGMEHLREVRAVTSANIQEFAGCKKIFGSLAFLEPSP 360
QY 375 DXVSIFYSLPLAP 387
DB 361 DGDPASNTAPLOP 373

RESULT 18

US-09-648-067A-1
; Sequence 1, Application US/09648067A
; Patent No. 6627196
; GENERAL INFORMATION:
; APPLICANT: Baughman, Sharon A.
; APPLICANT: Shak Steven
; TITLE OF INVENTION: Dosages for Treatment with Anti-ErbB2 Antibodies
; FILE REFERENCE: P1775R1
; CURRENT APPLICATION NUMBER: US/09/648,067A
; CURRENT FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/151,018
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: US 60/213,822
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 1
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-648-067A-1
Query Match 38.3%; Score 860; DB 4; Length 166;
Best Local Similarity 100.0%; Pred. No. 1.3e-68;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26 CTGTDMKRLRPASPETHLDMLRHLYQGCVVQGNLELTYLPTNASLSFLQDIOEVQGYVL 85
DB 1 CTGTDMKRLRPASPETHLDMLRHLYQGCVVQGNLELTYLPTNASLSFLQDIOEVQGYVL 60
QY 86 IAHNOVRQVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTPVTGASPGGLRELO 145
DB 61 IAHNOVRQVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTPVTGASPGGLRELO 120
QY 146 TEILKGGVLIQRNPOLCYQDTILWKDIFHKNQALTLIDTNRSPA 191
DB 121 TEILKGGVLIQRNPOLCYQDTILWKDIFHKNQALTLIDTNRSPA 166

RESULT 19

US-08-336-708A-9
 ; Sequence 9, Application US/08336708A
 ; Patent No. 5521295
 ; GENERAL INFORMATION:
 ; APPLICANT: Pacifici, Robert E.
 ; APPLICANT: Thomason, Arlen R.
 ; APPLICANT: Chang, Ming-Shi
 ; TITLE OF INVENTION: Hybrid Receptor Molecules
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Amgen Inc.
 ; STREET: 1840 Denavilland Drive
 ; CITY: Thousand Oaks
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 91320-1789
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/336,708A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Oleski, Nancy
 ; REFERENCE/DOCKET NUMBER: A-241A
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 644 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-336-708A-9

Query Match 35.3%; Score 793; DB 1; Length 644;

Best Local Similarity 45.3%; Pred. No. 6.4e-62;
 Matches 151; Conservative 50; Mismatches 118; Indels 14; Gaps 4;

```

QY 11 LLLALLPPGAA--STQVCTGTDKRLPASPETHLDMLRHLHYGGQVQGNLELTYPN 68
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 14 LLAALCPASRALEEKVKCGQTSNKLTLQGTFFEDHFLSLQRMFNCEVLGNLEITVQRN 73

QY 69 ASLSFLQDIOEVQGVYLLAHQVRQVPLQRLRIVRGTLQFEDNYALAVLDNGDPLNNTTP 128
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 74 YDLSFLKTIQEVAGVYLLAHNTVERPLENQLIIRGNMYENSALAVLSNYD----- 126

QY 129 VTGASPGGLRELQLRSLTEILKGVLIQRNPOLCYQDTILWKDIFHKNNQALTLIDNTR 188
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 244 RESDCLVCRKFRDEATCKDTCPPLMYNPTTYQMVNPEGRYTFGASCVTACPYNYLSTD 307

QY 308 VGSCTLVCPILNQEVTAEDGTQRCCKSPCAR 340
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 304 HGSCVRACGADSYEM-EEDGVRKCKCKEGPCR 335
  
```

RESULT 20

US-08-484-438-7
 ; Sequence 7, Application US/08484438
 ; Patent No. 5811098
 ; Patent No. 5811098 5780031
 ; GENERAL INFORMATION:

APPLICANT: Plowman, Gregory D.
 APPLICANT: Culouscou, Jean-Michel
 APPLICANT: Shovab, Mohammed
 APPLICANT: Siegall, Clay B.
 APPLICANT: Hellstr m, Ingegerd
 APPLICANT: Hellstr m, Karl E.
 TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,438
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/323,442
 FILING DATE: 14-OCT-1994
 APPLICATION NUMBER: US 08/150,704
 FILING DATE: 10-NOV-1993
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/981,165
 FILING DATE: 24-NOV-1992
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Mistrock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 5624-230
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1210 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-484-438-7

Query Match 35.3%; Score 793; DB 2; Length 1210;

Best Local Similarity 45.3%; Pred. No. 1.4e-61;
 Matches 151; Conservative 50; Mismatches 118; Indels 14; Gaps 4;

```

QY 11 LLLALLPPGAA--STQVCTGTDKRLPASPETHLDMLRHLHYGGQVQGNLELTYPN 68
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 14 LLAALCPASRALEEKVKCGQTSNKLTLQGTFFEDHFLSLQRMFNCEVLGNLEITVQRN 73

QY 69 ASLSFLQDIOEVQGVYLLAHQVRQVPLQRLRIVRGTLQFEDNYALAVLDNGDPLNNTTP 128
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 74 YDLSFLKTIQEVAGVYLLAHNTVERPLENQLIIRGNMYENSALAVLSNYD----- 126

QY 129 VTGASPGGLRELQLRSLTEILKGVLIQRNPOLCYQDTILWKDIFHKNNQALTLIDNTR 188
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 127 --ANKTGLKELPMENLQELHGAVRFSNNPACNVEISQWRDIVSSDFLSNMSPDQNH 183

QY 189 SRACHPCSPCKGSRGSGESSEDCQSLTRTVCAAGCA-RCKGPLPTDCCHQCAAGCTGP 247
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 184 LGSCQKCDPSCPNCSGAGGECNCKLTKIICAQCCSGRCGRKSPSCDCHNQCAAGCTGP 243

QY 248 KHSDDLACLHFNHSGICELHCPALVTYNTDTPESMNPBGRYTFGASCVTACPYNYLSTD 307
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  
```

Db 244 RESDCLVCRKFRDENTCKDTCPPLMLYNPTTYQMDVNPBGKYSFGATCVKCKPCPNYYVTD 303
QY 308 VGSCTLVCPLENQEVTAEDGTQRCCKSKPCAR 340
Db 304 HGSCVTRACGADSYEM-EDGVKCKCKCEGPRK 335
RESULT 21
US-08-475-035-4
; Sequence 4, Application US/08475035
; Patent No. 5985553
; GENERAL INFORMATION:
; APPLICANT: KING, C. R.
; APPLICANT: KRAUS, MATTHIAS H.
; APPLICANT: AARONSON, STUART A.
; TITLE OF INVENTION: HUKAN GENE RELATED TO BUT DISTINCT FROM
; TITLE OF INVENTION: EGF RECEPTOR GENE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; STREET: Suite 1200, 127 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,035
; FILING DATE: 7 Jun 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Peirymann, David G.
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.656
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1210 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-475-035-4
Query Match 35.3%; Score 793; DB 2; Length 1210;
Best Local Similarity 45.3%; Pred. No. 1.4e-61;
Matches 151; Conservative 50; Mismatches 118; Indels 14; Gaps 4;
QY 11 LLLALLPPGAA--STQVCTGTDMLKRLPASPEHLDMLRHLYQGVQVQGNLEITYLPTN 68
Db 14 LLAALCPASRALEKKVKQGSNKLTLQGTGTFDFHFLSLQRMFNCEVVLGNLEITYVORN 73
QY 59 ASLSFLQDIOEVQGVYLIHNRQVPLQRLRIVRGTLQFEDNYALAVLDNGDPLNNTTP 128
Db 74 YDLSFLKTIQEVAGYFLIALNTVERIPLENQIIRGNMYNSVALVLSYD----- 126
QY 129 VTGASPGGLREQLRLSLTEILKGGVLIQRPOLCYQDTILWKDIPIHKNQALALTIDNR 198
Db 127 ---ANKTKLKLPMENLQELILHGAVRFSNNPALCNVESIQWRDIVSSDFLSNMSDFQNH 183
QY 189 SRACHPCSPCKSGSCWGESSEDQSLRTVTCAGCA--RCKGPLPTDCHCQCAAGCTGP 247
Db 184 LGSQCKDPCSPCNGSCWAGGENSECQKLTIKIAQQCSGRCKRSPSCDCHNQCAGCTGP 243
QY 248 KHSDDLCLAHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYSTD 307
Db 244 RESDCLVCRKFRDENTCKDTCPPLMLYNPTTYQMDVNPBGKYSFGATCVKCKPCPNYYVTD 303

QY 308 VGSCTLVCPLENQEVTAEDGTQRCCKSKPCAR 340
Db 304 HGSCVTRACGADSYEM-EDGVKCKCKCEGPRK 335
RESULT 22
US-09-570-454-2
; Sequence 2, Application US/09570454
; Patent No. 6395743
; GENERAL INFORMATION:
; APPLICANT: Department of Veterans Affairs
; TITLE OF INVENTION: Isolation and characterization of epidermal growth
; TITLE OF INVENTION: factor related protein
; FILE REFERENCE: 107999.00106
; CURRENT APPLICATION NUMBER: US/09/570,454
; CURRENT FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: 60/134,200
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-570-454-2
Query Match 34.4%; Score 773; DB 4; Length 478;
Best Local Similarity 45.7%; Pred. No. 2.6e-60;
Matches 154; Conservative 43; Mismatches 118; Indels 22; Gaps 6;
QY 11 LLLALLPPGAA--STQVCTGTDMLKRLPASPEHLDMLRHLYQGVQVQGNLEITYLPTN 68
Db 14 LLLTALCAAGGALREKVKVQGSNRLTLQGTGTFDFHFLSLQRMFNCEVVLGNLEITYVORN 73
QY 69 ASLSFLQDIOEVQGVYLIHNRQVPLQRLRIVRGTLQFEDNYALAVLDNGDPLNNTTP 128
Db 74 YDLSFLKTIQEVAGYFLIALNTVERIPSEDQIIRGNALYENTYALAILSN----- 124
QY 129 VTGASPGGLREQLRLSLTEILKGGVLIQRPOLCYQDTILWKDIPIHKNQALALTIDNR 184
Db 125 -YGNRTGLRELPENLQELILHGAVRFSNNPALCNMTTQWRDIVQNVFMSMDL--- 180
QY 185 DTNRSRACHPCSPCKSGSCWGESSEDQSLRTVTCAGCA--RCKGPLPTDCHCQCAAG 243
Db 181 -QSHPSCKPCDFSPCNGSCWGGENSECQKLTIKIAQQCSGRCKRSPSCDCHNQCAG 239
QY 244 CTGPKGSDCLAHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNY 303
Db 240 CTGPKGSDCLVCPQDDEATCKDTCPPLMLYNPTTYQMDVNPBGKYSFGATCVKCPNRY 299
QY 304 LSTDVGSCTLVCPLENQEVTAEDGTQRCCKSKPCAR 340
Db 300 VVTDHGSCTVACGPDYEV-EEGIRKCKCKDGPRK 335

RESULT 23
US-09-867-521-2
; Sequence 2, Application US/09867521
; Patent No. 6582934
; GENERAL INFORMATION:
; APPLICANT: Department of Veterans Affairs
; TITLE OF INVENTION: Isolation and characterization of epidermal growth
; TITLE OF INVENTION: factor related protein
; FILE REFERENCE: 111828-00103
; CURRENT APPLICATION NUMBER: US/09/867,521
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/134,200
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/570,454
; PRIOR FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2

```
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-867-521-2

Query Match      34.4%; Score 773; DB 4; Length 478;
Best Local Similarity 45.7%; Pred. No. 2.6e-60;
Matches 154; Conservative 43; Mismatches 118; Indels 22; Gaps 6;

QY 11 LLLALLPPGAA--STQVCTGTDMLRLPASPEHLDMLRHLYOGCOVVOGNLELTVLPTN 68
Db 14 LUTALCAAGALEEKVCQGTNSRLTQLGTFEDHFLSLQRMVYNNCEVVLGNLEITVQRN 73
QY 69 ASLSPLQDIQEVQGVYLIHNVQVPLQRLRIVRGTLQFEDNYALAVLDNGDPLNTPP 128
Db 74 YDLSPLKTIQEVAGYELIALNTVERIPSEDQIIRGNALYENTYALAILSN----- 124
QY 129 VTGASPGGIRELQRLSLTEILKGGVLIQNPOLCYQDTILWKDI----EKKNNQLALTIL 184
Db 125 -YGTNTGLRELPMRLQELIGAVFNSNPILCNMDTIQMRDIVQNVFWSNMSMDL--- 180
QY 185 DTRNSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCAAGCA-RCKGPLEPTDCCHBQCAAG 243
Db 181 -QSHPPSCPKDPCSPCNGSGCWGGEENCKVLTKIICAQCSHRCRGRSPSDCCHNQCAAG 239
QY 244 CTGPKHSDCLACHFNHSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACPNY 303
Db 240 CTGPKHSDCLVQKQFQDEATCKDTCPPLMLYNTTYQMDNPEGKYSFGATCVKCPRY 299
QY 304 LSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR 340
Db 300 VVTDHSGCVRACGPDYEV-EEDGIRKCKKCDPCPK 335

RESULT 24
US-08-484-438-10
; Sequence 10, Application US/08484438
; Patent No. 5811098
; GENERAL INFORMATION:
; APPLICANT: Florman, Gregory D.
; APPLICANT: Culouscou, Jean-Michel
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Siegall, Clay B.
; APPLICANT: Hellstr m, Ingegerd
; APPLICANT: Hellstr m, Karl E
; TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,438
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/323,442
; FILING DATE: 14-OCT-1994
; APPLICATION NUMBER: US 08/150,704
; FILING DATE: 10-NOV-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/981,165
; FILING DATE: 24-NOV-1992
```

```
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: S624-230
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 911 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-484-438-10

Query Match      34.2%; Score 768.5; DB 2; Length 911;
Best Local Similarity 45.4%; Pred. No. 1.5e-59;
Matches 153; Conservative 44; Mismatches 123; Indels 17; Gaps 6;

QY 9 WGLLLALLPPGAA---STQVCTGTDMLRLPASPEHLDMLRHLYOGCOVVOGNLELTY 64
Db 8 WWWVLLVAAGTVQPSDSQSVCACTENKLSLSLDEQQYRALRKYENCEVVMGNLEITS 67
QY 65 LPTNASLSFLQDIQEVQGVYLIHNVQVPLQRLRIVRGTLQFEDNYALAVLDNGDPLN 124
Db 68 IEHRDLSFLRSVREVTGYLVALNQFRLPLENRIIRGTNKLVDYALALFLNRYKDG 127
QY 125 NTPPTVGASPGGIRELQRLSLTEILKGGVLIQNPOLCYQDTILWKDIFKNNQLALTIL 184
Db 128 NF-----GLQELGLKNTLILNGGVYVDQNKFLCYADTIHQDILVRNFWPSNLTIV 178
QY 185 DTRNSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCAAGC-ARCKGPLEPTDCCHBQCAAG 243
Db 179 STNGSSGGCRCHKSCGTG-RCWGPTENHCQTLTRTVCAEQCDRCYGVVSDCHRECAAG 237
QY 244 CTGPKHSDCLACHFNHSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACPNY 303
Db 238 CSGPADTDFACMFNDSGACVTCQPTFTVNTPTFFQLEHFNKITYTGACFCVKKCPHNF 297
QY 304 LSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR 340
Db 298 V-VDSSSCVRACPSKMEV-EENGKMKCKPCTDICK 332

RESULT 25
US-08-484-438-4
; Sequence 4, Application US/08484438
; Patent No. 5811098
; GENERAL INFORMATION:
; APPLICANT: Florman, Gregory D.
; APPLICANT: Culouscou, Jean-Michel
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Siegall, Clay B.
; APPLICANT: Hellstr m, Ingegerd
; APPLICANT: Hellstr m, Karl E
; TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
```


Sequence 4, Application US/07978895
Patent No. 5480968
GENERAL INFORMATION:
APPLICANT: Kraus, Matthias H.
APPLICANT: Aaronson, Stuart A.
TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR RECEPTOR, ANTIGEN THEREO, AND
TITLE OF INVENTION: BIOASSAYS AND METHODS RELATED THEREO
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Suite 400
STREET: 133 Carnegie Way, N.W.
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.A.
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/978,895
FILING DATE: 1992/11/10
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/444,406
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Perryman, David G.
REGISTRATION NUMBER: 33,438
REFERENCE/DOCKET NUMBER: 1414-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 688-0770
TELEFAX: (404) 688-9880
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1342 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-978-895-4

Query Match 32.7%; Score 734; DB 1; Length 1342;
Best Local Similarity 44.0%; Pred. No. 2.8e-56;
Matches 147; Conservative 49; Mismatches 120; Indels 18; Gaps 8;

QY 10 GLLALLPPGAA--STQVCTGTDMLKRLPASPETHDMLRHLVYQGVVQGNLELYLPT 67
DB 11 GLLFSLARGSEVGNQAVCPGTLNGLSVTGDENQVQTLKLYRCEVVMGNLEIVLTGH 70
QY 68 NASLSFLQDIOEVQGVYLIHNVQVPLQRLRVIRVGTQLFEDNYALVLDNGDPLNNTT 127
DB 71 NADLSFLQWIREVTGYLVAMNEFTLPNLRVVRGTQVYDGKFAIFVM----LNNT 125
QY 128 PVTGASPGGLRELQLRSLTEILKGGVLIQRNPQLCVQDTILWKDIFHKNQALALTLDTN 187
DB 126 ----NSSHALRQLRLTQLTEILSGGVYIEKNDKLCMDTIDWRDVRD---AEIVVKD 178
QY 188 RSRACHPCSPMKGSRGWESSEDCQSLTRTVACGGC-ARCKGPIPTDCCHQCAAGCTG 246

QY 10 GLLALLPPGAA--STQVCTGTDMLKRLPASPETHDMLRHLVYQGVVQGNLELYLPT 67
DB 11 GLLFSLARGSEVGNQAVCPGTLNGLSVTGDENQVQTLKLYRCEVVMGNLEIVLTGH 70
QY 68 NASLSFLQDIOEVQGVYLIHNVQVPLQRLRVIRVGTQLFEDNYALVLDNGDPLNNTT 127
DB 71 NADLSFLQWIREVTGYLVAMNEFTLPNLRVVRGTQVYDGKFAIFVM----LNNT 125
QY 128 PVTGASPGGLRELQLRSLTEILKGGVLIQRNPQLCVQDTILWKDIFHKNQALALTLDTN 187
DB 126 ----NSSHALRQLRLTQLTEILSGGVYIEKNDKLCMDTIDWRDVRD---AEIVVKD 178
QY 188 RSRACHPCSPMKGSRGWESSEDCQSLTRTVACGGC-ARCKGPIPTDCCHQCAAGCTG 246

RESULT 28

US-08-484-438-9
Sequence 9, Application US/08484438
Patent No. 5811098
GENERAL INFORMATION:
APPLICANT: Plowman, Gregory D.
APPLICANT: Culouscou, Jean-Michel
APPLICANT: Shoyab, Mohammed
APPLICANT: Siegall, Clay B.
APPLICANT: Hellstr m, Ingegerd
APPLICANT: Hellstr m, Karl E.
TITLE OF INVENTION: HERA HUMAN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,438
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,442
FILING DATE: 14-OCT-1994
APPLICATION NUMBER: US 08/150,704
FILING DATE: 10-NOV-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/981,165
FILING DATE: 24-NOV-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leelle
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1342 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-438-9

Query Match 32.7%; Score 734; DB 2; Length 1342;
Best Local Similarity 44.0%; Pred. No. 2.8e-56;
Matches 147; Conservative 49; Mismatches 120; Indels 18; Gaps 8;

QY 10 GLLALLPPGAA--STQVCTGTDMLKRLPASPETHDMLRHLVYQGVVQGNLELYLPT 67
DB 11 GLLFSLARGSEVGNQAVCPGTLNGLSVTGDENQVQTLKLYRCEVVMGNLEIVLTGH 70
QY 68 NASLSFLQDIOEVQGVYLIHNVQVPLQRLRVIRVGTQLFEDNYALVLDNGDPLNNTT 127
DB 71 NADLSFLQWIREVTGYLVAMNEFTLPNLRVVRGTQVYDGKFAIFVM----LNNT 125
QY 128 PVTGASPGGLRELQLRSLTEILKGGVLIQRNPQLCVQDTILWKDIFHKNQALALTLDTN 187
DB 126 ----NSSHALRQLRLTQLTEILSGGVYIEKNDKLCMDTIDWRDVRD---AEIVVKD 178
QY 188 RSRACHPCSPMKGSRGWESSEDCQSLTRTVACGGC-ARCKGPIPTDCCHQCAAGCTG 246

Db 179 NGRSCPPCHEVCKG-RCWGPGECCQTLTKTICAPQCNHGFCEGPNPQCCHDECAGGCSG 237
Qy 247 PKHSDCLAHNHSIGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLST 306
Db 238 PQDIDCFACRHFNDGACVPCPCPLVYNKLTFLQEPNPHTKYQYGVGVASCPHNFV-V 296
Qy 307 DVGSCITVCLPHNQEVTAEDGTQRCCKSKPCAR 340
Db 297 DQTSVCRACPPDKMEVD-KNGLKWCPEOGGLCPK 329

RESULT 29
US-08-473-119-4
; Sequence 4, Application US/08473119
; Patent No. 5820859
; GENERAL INFORMATION:
; APPLICANT: Kraus, Matthias H.
; APPLICANT: Aaronson, Stuart A.
; TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
; TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR RECEPTOR, ANTIGEN THEREOF, AND
; TITLE OF INVENTION: BIOASSAYS AND METHODS RELATED THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Suite 400
; STREET: 133 Carnegie Way, N.W.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.A.
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/473,119
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/978,895
; FILING DATE: 10-NOV-1992
; APPLICATION NUMBER: US 07/444,406
; FILING DATE: 01-DEC-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Perryman, David G.
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414-028
; TELEPHONE: (404) 688-0770
; TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-473-119-4

Query Match 32.7%; Score 734; DB 2; Length 1342;
Best Local Similarity 44.0%; Pred. No. 2.8e-56;
Matches 147; Conservative 49; Mismatches 120; Indels 18; Gaps 8;

Qy 10 GLLALLPPGAA--STQVCTGTDMLKRLPASPETHLDMLRHLYGCGVQVQGNLELYLPT 67
Db 11 GLLFLARGSEGVNSQAVCPGTLNGLSVTGAENQYQTLKLYERCEVVMGNLELYLPT 70
Qy 68 NASLSFLQDIQEVGVVLIHNVQVPLQRLRVIRVGTQLFEDNVAVLDNGDPLNNTT 127
Db 71 NADLSFLQIREVTVGVVLIHNVQVPLQRLRVIRVGTQVYDGFALFVW-----LNYNT 125
Qy 128 PVTGASPGGLRELQFSLTEILKGGVLIQRPOLCYQDILMKDIFPKNNALFLIDTN 187

Db 126 ----NSSHALRQLRLTQLTEILSGGVYIENKDKLCHMDTIDMRDIVDRD---AEIVVKD 178
Qy 188 RSRACPCSPMKGSRGWGSSSDQSLRTVCAGGC-ARCKGGLPTDCCHEQCAAGCTG 246
Db 179 NGRSCPPCHEVCKG-RCWGPGECCQTLTKTICAPQCNHGFCEGPNPQCCHDECAGGCSG 237
Qy 247 PKHSDCLAHNHSIGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLST 306
Db 238 PQDIDCFACRHFNDGACVPCPCPLVYNKLTFLQEPNPHTKYQYGVGVASCPHNFV-V 296
Qy 307 DVGSCITVCLPHNQEVTAEDGTQRCCKSKPCAR 340
Db 297 DQTSVCRACPPDKMEVD-KNGLKWCPEOGGLCPK 329

RESULT 30
US-08-475-352-4
; Sequence 4, Application US/08475352
; Patent No. 5916755
; GENERAL INFORMATION:
; APPLICANT: Kraus, Matthias H.
; APPLICANT: Aaronson, Stuart A.
; TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
; TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR RECEPTOR, ANTIGEN THEREOF, AND
; TITLE OF INVENTION: BIOASSAYS AND METHODS RELATED THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Suite 400
; STREET: 133 Carnegie Way, N.W.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.A.
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,352
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/978,895
; FILING DATE:
; APPLICATION NUMBER: US 07/444,406
; FILING DATE: 01-DEC-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Perryman, David G.
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414-028
; TELEPHONE: (404) 688-0770
; TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-475-352-4

Query Match 32.7%; Score 734; DB 2; Length 1342;
Best Local Similarity 44.0%; Pred. No. 2.8e-56;
Matches 147; Conservative 49; Mismatches 120; Indels 18; Gaps 8;

Qy 10 GLLALLPPGAA--STQVCTGTDMLKRLPASPETHLDMLRHLYGCGVQVQGNLELYLPT 67
Db 11 GLLFLARGSEGVNSQAVCPGTLNGLSVTGAENQYQTLKLYERCEVVMGNLELYLPT 70
Qy 68 NASLSFLQDIQEVGVVLIHNVQVPLQRLRVIRVGTQLFEDNVAVLDNGDPLNNTT 127
Db 71 NADLSFLQIREVTVGVVLIHNVQVPLQRLRVIRVGTQVYDGFALFVW-----LNYNT 125

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2004, 04:08:53 ; Search time 85.8193 Seconds
(without alignments)
1519.797 Million cell upda

Title: US-09-506-079H-2

Perfect score: 2247

Sequence: 1 MELAALCRWGLLLALLPPGA.....VGRGXDPDAHVAVXLSRYEG 419

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283816 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq Length: 0

Minimum	DB seq	length:	0
Maximum	DB seq	length:	2000000000

Post-processing: Minimum Match 0%

POST-PROCESSING: MINIMUM MATCH 0%
Maximum Match 100%

Maximum Match 100%
Listing first 90 summaries

Database : Published Applications AA:★

1: /cgm2_6/prodata/2/pubpaa/US07 PUBCOMB.pcp:
2: /cgm2_6/prodata/2/pubpaa/PCT NEW_PUB.pcp:
3: /cgm2_6/prodata/2/pubpaa/US06 NEW_PUB.pcp:
4: /cgm2_6/prodata/2/pubpaa/US06 PUBCOMB.pcp:
5: /cgm2_6/prodata/2/pubpaa/US07 NEW_PUB.pcp:
6: /cgm2_6/prodata/2/pubpaa/PCTUS PUBCOMB.pcp:
7: /cgm2_6/prodata/2/pubpaa/US08 NEW_PUB.pcp:
8: /cgm2_6/prodata/2/pubpaa/US08 PUBCOMB.pcp:
9: /cgm2_6/prodata/2/pubpaa/US09A PUBCOMB.pcp:
10: /cgm2_6/prodata/2/pubpaa/US09B PUBCOMB.pcp:
11: /cgm2_6/prodata/2/pubpaa/US09C PUBCOMB.pcp:
12: /cgm2_6/prodata/2/pubpaa/US09 NEW_PUB.pcp:
13: /cgm2_6/prodata/2/pubpaa/US10A PUBCOMB.pcp:
14: /cgm2_6/prodata/2/pubpaa/US10C PUBCOMB.pcp:
15: /cgm2_6/prodata/2/pubpaa/US10 PUBCOMB.pcp:
16: /cgm2_6/prodata/2/pubpaa/US10 NEW_PUB.pcp:
17: /cgm2_6/prodata/2/pubpaa/US60 NEW_PUB.pcp:
18: /cgm2_6/prodata/2/pubpaa/US60 PUBCOMB.pcp:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2190	97.5	419	12	US-10-344-470-2		Sequence 2, Appli
2	2178	96.9	419	16	US-10-302-663-2		Sequence 2, Appli
3	1874.5	83.4	720	15	US-10-412-804A-6		Sequence 6, Appli
4	1873.5	83.4	645	9	US-09-921-161-1		Sequence 1, Appli
5	1873.5	83.4	645	15	US-10-268-501-13		Sequence 13, Appli
6	1873.5	83.4	645	14	US-10-608-626-13		Sequence 13, Appli
7	1873.5	83.4	653	9	US-09-854-356-3		Sequence 3, Appli
8	1873.5	83.4	685	15	US-10-412-804A-4		Sequence 4, Appli
9	1873.5	83.4	690	15	US-10-412-804A-11		Sequence 11, Appli
10	1873.5	83.4	712	9	US-09-854-356-7		Sequence 7, Appli
11	1873.5	83.4	715	15	US-10-412-804A-10		Sequence 10, Appli
12	1873.5	83.4	919	9	US-09-854-356-6		Sequence 6, Appli
13	1873.5	83.4	1253	14	US-10-146-473-72		Sequence 72, Appli
14	1873.5	83.4	1355	9	US-09-811-123-9		Sequence 9, Appli
15	1873.5	83.4	1355	9	US-09-811-115-3		Sequence 3, Appli

89 263.5 11.7 1382 16 US-10-408-7658A-149 Sequence 149, App
90 263 11.7 1297 12 US-10-333-314-1 Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-344-470-2
; Sequence 2, Application US/10344470
; Publication No. US20040052796A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; TITLE OF INVENTION: EXPRESSION OF HERSTATIN, AN ALTERNATIVE TO HER-2/NEU PRODUCT, IN
; TITLE OF INVENTION: THAT EXPRESS EITHER P185HER-2 OR THE EGF RECEPTOR INHIBITS RECH
; TITLE OF INVENTION: CELL GROWTH
; FILE REFERENCE: 49321-81
; CURRENT APPLICATION NUMBER: US/10/344,470
; PRIOR FILING DATE: 2003-06-09
; PRIOR APPLICATION NUMBER: PCT / US01/25502
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 09/638,834
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIt version 3.1
; SEQ ID NO 2
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (342)..(342)
; OTHER INFORMATION: Applicants herein disclose Thr and Ser sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (345)..(345)
; OTHER INFORMATION: Applicants herein disclose Leu and Pro sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (346)..(346)
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (356)..(356)
; OTHER INFORMATION: Applicants herein disclose Leu and Gln sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (358)..(358)
; OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (361)..(361)
; OTHER INFORMATION: Applicants herein disclose Gly, Asp, Ala and Val sequence variant
; OTHER INFORMATION: \$ at this position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (376)..(376)
; OTHER INFORMATION: Applicants herein disclose Leu and Ile sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (394)..(394)
; OTHER INFORMATION: Applicants herein disclose Pro and Arg sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (404)..(404)
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this

; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (413)..(413)
; OTHER INFORMATION: Applicants herein disclose Asp and Asn sequence variants at thi
; OTHER INFORMATION: position
US-10-344-470-2

Query Match 97.5%; Score 2190; DB 12; Length 419;
Best Local Similarity 98.6%; Pred. No. 2.9e-182;
Matches 413; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLALIPGGAATVCTGTDMLRLPASPETHLDMLRLHLYQGCVVQGNL 60
DB 1 MELAALCRWGLLALLPFGAATVCTGTDCKLRUPASPETHLDMLRLHLYQGCVVQGNL 60

QY 61 ELTYLPTNASLSFLODIQEVQGYVLIHNNQVQPLQRLRIVRGTQLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLODIQEVQGYVLIHNNQVQPLQRLRIVRGTQLFEDNYALAVLDNG 120

QY 121 DPLNNTTPTVGASPGGLRELQRLSLTEILKGGVLIQRPQLCYQDTILMKDIFHKNNQLA 180
DB 121 DPLNNTTPTVGASPGGLRELQRLSLTEILKGGVLIQRPQLCYQDTILMKDIFHKNNQLA 180

QY 181 LTLIDTNRGRACHPCSPCKGSRGWSSESSEDQSLTRTVGAGGCARCKGPLEPTDCCHQC 240
DB 181 LTLIDTNRGRACHPCSPCKGSRGWSSESSEDQSLTRTVGAGGCARCKGPLEPTDCCHQC 240

QY 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTLVCPHNNQVTAEDGTQRCBCKSKPCARGKHSXPRPAAVPPVPRXQP 360
DB 301 YNKLSTDVGSCTLVCPHNNQVTAEDGTQRCBCKSKPCARGKHSXPRPAAVPPVPRXQP 360

QY 361 XPAHPVLSPLRPSWDXVSIFYSLPLAPLDPTSVXISPVSGWGRGXDPAHVAVKLSRYEG 419
DB 361 XPAHPVLSPLRPSWDXVSIFYSLPLAPLDPTSVXISPVSGWGRGXDPAHVAVKLSRYEG 419

RESULT 2
US-10-302-663-2
; Sequence 2, Application US/10302663
; Publication No. US20040022785A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; TITLE OF INVENTION: EXPRESSION OF HERSTATIN, AN ALTERNATIVE HER-2/NEU PRODUCT, IN C
; TITLE OF INVENTION: EXPRESS EITHER P185HER-2 OR THE EGF RECEPTOR INHIBITS RECEPTOR
; TITLE OF INVENTION: GROWTH
; FILE REFERENCE: 49321-73
; CURRENT APPLICATION NUMBER: US/10/302,663
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 09/638,834
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 2
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 342
; OTHER INFORMATION: Applicants herein disclose Thr and Ser sequence variants at thi
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 345
; OTHER INFORMATION: Applicants herein disclose Leu and Pro sequence variants at thi
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 346
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at thi
; FEATURE:

```
; NAME/KEY: VARIANT
; LOCATION: 356
; OTHER INFORMATION: Applicants herein disclose Leu and Gln sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 356
; OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 361
; OTHER INFORMATION: Applicants herein disclose Gly, Asp, Ala and Val sequence variants
; OTHER INFORMATION: Position
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 376
; OTHER INFORMATION: Applicants herein disclose Leu and Ile sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 394
; OTHER INFORMATION: Applicants herein disclose Pro and Arg sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 404
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 413
; OTHER INFORMATION: Applicants herein disclose Asp and Asn sequence variants at this
; OTHER INFORMATION: Applicants herein disclose Asp and Asn sequence variants at this
US-10-302-663-2

Query Match          96.9%; Score 2178; DB 16; Length 419;
Best Local Similarity 98.1%; Pred. No. 3.2e-181;
Matches 411; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQGCQVVGQNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQGCQVVGQNL 60

QY 61 ELTYLPTNASLSFLQDIQVQGVLIAHNQVQVPLQRLIRIVRGTLQEDNVALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQVQGVLIAHNQVQVPLQRLIRIVRGTLQEDNVALAVLNG 120

QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQVQVPLQRLIRIVRGTLQEDNVALAVLNG 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQVQVPLQRLIRIVRGTLQEDNVALAVLNG 180

QY 181 LTLIDNTRSRACHPCSPCKGSRGWESSEDQSLTRTVAGGCARCKGPLPTDCHEQC 240
DB 181 LTLIDNTRSRACHPCSPCKGSRGWESSEDQSLTRTVAGGCARCKGPLPTDCHEQC 240

QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESCPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESCPNPEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTLVCPLENQVTAEDGTQRCCKSPKPCAR---GXHSXXPRPAAVVPXRP 360
DB 301 YNYLSTDVGSCTLVCPLENQVTAEDGTQRCCKSPKPCAR---GXHSXXPRPAAVVPXRP 360

QY 361 XPAHPVLSLRPSWDVSAFYSPLPLAPLPTSVXISPVSVGKXGDPDAHVAVXLSRYEG 419
DB 361 XPAHPVLSLRPSWDVSAFYSPLPLAPLPTSVXISPVSVGKXGDPDAHVAVXLSRYEG 419

RESULT 3
US-10-412-804A-6
; Sequence 6, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Jing, Shuqian
; APPLICANT: Tatarewicz, Suzanna
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
; FILE REFERENCE: 01-1624-A
```

```
; CURRENT APPLICATION NUMBER: US/10/412,804A
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,912
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-412-804A-6

Query Match          83.4%; Score 1874.5; DB 15; Length 720;
Best Local Similarity 87.9%; Pred. No. 1.8e-154;
Matches 356; Conservative 3; Mismatches 31; Indels 15; Gaps 4;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQGCQVVGQNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQGCQVVGQNL 60

QY 61 ELTYLPTNASLSFLQDIQVQGVLIAHNQVQVPLQRLIRIVRGTLQEDNVALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQVQGVLIAHNQVQVPLQRLIRIVRGTLQEDNVALAVLNG 120

QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQVQVPLQRLIRIVRGTLQEDNVALAVLNG 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQVQVPLQRLIRIVRGTLQEDNVALAVLNG 180

QY 181 LTLIDNTRSRACHPCSPCKGSRGWESSEDQSLTRTVAGGCARCKGPLPTDCHEQC 240
DB 181 LTLIDNTRSRACHPCSPCKGSRGWESSEDQSLTRTVAGGCARCKGPLPTDCHEQC 240

QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESCPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESCPNPEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTLVCPLENQVTAEDGTQRCCKSPKPCAR---GXHSXXPRPAAVVPXRP 357
DB 301 YNYLSTDVGSCTLVCPLENQVTAEDGTQRCCKSPKPCAR---GXHSXXPRPAAVVPXRP 357

QY 358 XQXPAA-----HPVLSFLRPSWDVSAFYSPLPLAPLPTSVXISPVSVGKXGDPDAHVAVXLSRYEG 390
DB 361 IQEFAGACKKIFGSLAFLEPSEFVGSLCQQAQGVQWYDGLGSLQLP 405

RESULT 4
US-09-921-161-1
; Sequence 1, Application US/099211161
; Patent No. US2002090662A1
; GENERAL INFORMATION:
; APPLICANT: Ralph, Peter
; TITLE OF INVENTION: ANALYTICAL METHOD
; FILE REFERENCE: GENENT.066A
; CURRENT APPLICATION NUMBER: US/09/921,161
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/225,433
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-921-161-1

Query Match          83.4%; Score 1873.5; DB 9; Length 645;
Best Local Similarity 89.3%; Pred. No. 1.9e-154;
Matches 352; Conservative 3; Mismatches 33; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQGCQVVGQNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQGCQVVGQNL 60
```

QY 61 ELTYLPTNASLFLQDIOEVGYVLIHNOVQVPLQRLIRVGTQLPEDNYALAVLDNG 120
DB 61 ELTYLPTNASLFLQDIOEVGYVLIHNOVQVPLQRLIRVGTQLPEDNYALAVLDNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGVLIQNPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGVLIQNPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPMCKGSRGWSSEDCQSLTRITVCAGGCARCKGPLEPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMCKGSRGWSSEDCQSLTRITVCAGGCARCKGPLEPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAPVPXR 357
DB 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAPVPXR 357
QY 358 XQXPXPA---HPVLSFLRPSWDKXSAFYSLPLAP 387
DB 361 IQBFAGCKKIFGSLAFPLPSFDGDPASNTAPLQP 394

RESULT 5
US-10-268-501-13
; Sequence 13, Application US/10268501
; Publication No. US20030086924A1
; GENERAL INFORMATION:
; APPLICANT: Sliwowski, Mark X.
; TITLE OF INVENTION: Treatment with Anti-ErbB2 Antibodies
; FILE REFERENCE: P146782P1
; CURRENT APPLICATION NUMBER: US/10/268,501
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 09/602,812
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141,316
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 13
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-268-501-13

Query Match 83.4%; Score 1873.5; DB 14; Length 645;
Best Local Similarity 89.3%; Pred. No. 1.9e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASSTVCTGDMKRLPASPETHLDMRLHLYQGCVVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASSTVCTGDMKRLPASPETHLDMRLHLYQGCVVQGNL 60
QY 61 ELTYLPTNASLFLQDIOEVGYVLIHNOVQVPLQRLIRVGTQLPEDNYALAVLDNG 120
DB 61 ELTYLPTNASLFLQDIOEVGYVLIHNOVQVPLQRLIRVGTQLPEDNYALAVLDNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGVLIQNPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGVLIQNPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPMCKGSRGWSSEDCQSLTRITVCAGGCARCKGPLEPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMCKGSRGWSSEDCQSLTRITVCAGGCARCKGPLEPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAPVPXR 357
DB 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAPVPXR 357

QY 358 XQXPXPA---HPVLSFLRPSWDKXSAFYSLPLAP 387
DB 361 IQBFAGCKKIFGSLAFPLPSFDGDPASNTAPLQP 394
RESULT 6
US-10-608-626-13
; Sequence 13, Application US/10608626
; Publication No. US20040013667A1
; GENERAL INFORMATION:
; APPLICANT: Kelsey, Stephen M.
; APPLICANT: Sliwowski, Mark X.
; TITLE OF INVENTION: Treatment with Anti-ErbB2 Antibodies
; FILE REFERENCE: P146782P2
; CURRENT APPLICATION NUMBER: US/10/608,626
; PRIOR FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: US 10/268,501
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 09/602,812
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141,316
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 13
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-608-626-13

Query Match 83.4%; Score 1873.5; DB 15; Length 645;
Best Local Similarity 89.3%; Pred. No. 1.9e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASSTVCTGDMKRLPASPETHLDMRLHLYQGCVVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASSTVCTGDMKRLPASPETHLDMRLHLYQGCVVQGNL 60
QY 61 ELTYLPTNASLFLQDIOEVGYVLIHNOVQVPLQRLIRVGTQLPEDNYALAVLDNG 120
DB 61 ELTYLPTNASLFLQDIOEVGYVLIHNOVQVPLQRLIRVGTQLPEDNYALAVLDNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGVLIQNPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGVLIQNPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPMCKGSRGWSSEDCQSLTRITVCAGGCARCKGPLEPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMCKGSRGWSSEDCQSLTRITVCAGGCARCKGPLEPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAPVPXR 357
DB 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAPVPXR 357
QY 358 XQXPXPA---HPVLSFLRPSWDKXSAFYSLPLAP 387
DB 361 IQBFAGCKKIFGSLAFPLPSFDGDPASNTAPLQP 394

RESULT 7
US-09-854-356-3
; Sequence 3, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins

```
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
US-09-854-356-3

Query Match      83.4%; Score 1873.5; DB 9; Length 653;
Best Local Similarity 89.3%; Pred. No. 2e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASSTOCTGDMKRLPASPETHLDMLRHLVGGCQVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASSTOCTGDMKRLPASPETHLDMLRHLVGGCQVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIAHNQVRQVPLQRLRIVRGTLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIAHNQVRQVPLQRLRIVRGTLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHNQOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHNQOLA 180
QY 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVAGGCARCKGPIPTDCCHQOC 240
DB 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVAGGCARCKGPIPTDCCHQOC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPLNQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAVVPVEXR 357
DB 301 YNYLSTDVGSCTLVCPLNQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAVVPVEXR 357
QY 358 XQPXPA----HPVLSFLRPSWDVSAFYSLPLAP 387
DB 361 IQEFAGCKKIFGSLAFLPESFDGDPASNTAPLOP 394

RESULT 9
US-10-412-804A-11
; Sequence 11, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Jing, Shuqian
; APPLICANT: Tatarewicz, Suzanna
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
; FILE REFERENCE: 01-1624-A
; CURRENT APPLICATION NUMBER: US/10/412,804A
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,912
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-412-804A-11

Query Match      83.4%; Score 1873.5; DB 15; Length 690;
Best Local Similarity 89.3%; Pred. No. 2.1e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASSTOCTGDMKRLPASPETHLDMLRHLVGGCQVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASSTOCTGDMKRLPASPETHLDMLRHLVGGCQVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIAHNQVRQVPLQRLRIVRGTLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIAHNQVRQVPLQRLRIVRGTLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHNQOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHNQOLA 180
QY 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVAGGCARCKGPIPTDCCHQOC 240
DB 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVAGGCARCKGPIPTDCCHQOC 240

US-09-854-356-3

Query Match      83.4%; Score 1873.5; DB 9; Length 653;
Best Local Similarity 89.3%; Pred. No. 2e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASSTOCTGDMKRLPASPETHLDMLRHLVGGCQVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASSTOCTGDMKRLPASPETHLDMLRHLVGGCQVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIAHNQVRQVPLQRLRIVRGTLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIAHNQVRQVPLQRLRIVRGTLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHNQOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHNQOLA 180
QY 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVAGGCARCKGPIPTDCCHQOC 240
DB 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVAGGCARCKGPIPTDCCHQOC 240

RESULT 8
US-10-412-804A-4
; Sequence 4, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Jing, Shuqian
; APPLICANT: Tatarewicz, Suzanna
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
; FILE REFERENCE: 01-1624-A
; CURRENT APPLICATION NUMBER: US/10/412,804A
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,912
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-412-804A-4

Query Match      83.4%; Score 1873.5; DB 15; Length 685;
```


Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db |||||
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db |||||
Qy 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 357
Db |||||
Qy 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 360
Db |||||
Qy 358 XQXPA----HPVLSFLRPSWDXVSFYSPLAP 387
Db |||||
Qy 361 IQBFAGCKKIFGSLAFPSFDPGDPASNTAPLQ 394
Db |||||

RESULT 10

US-09-854-356-7
; Sequence 7, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: SmithKline Beecham Biologicals S. A.
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 7
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein
; OTHER INFORMATION: of ECD and delta PD of human HER-2/neu
US-09-854-356-7

Query Match 83.4%; Score 1873.5; DB 9; Length 712;

Best Local Similarity 89.3%; Pred. No. 2.2e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

Qy 1 MELAALCRWGLLLALLPPGAASCTVCTGDMKLRPLASPETHLDMLRHLYQGCVVQGNL 60
Db |||||
Qy 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 120
Db |||||
Qy 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 120
Db |||||
Qy 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGGVLIQNPOLCYQDTILWKDIFHKNOLA 180
Db |||||
Qy 181 LTLDITNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCCAGGCARCKGPLEPTDCCHEQC 240
Db |||||
Qy 181 LTLDITNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCCAGGCARCKGPLEPTDCCHEQC 240
Db |||||
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db |||||
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db |||||
Qy 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 357
Db |||||
Qy 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 360
Db |||||
Qy 358 XQXPA----HPVLSFLRPSWDXVSFYSPLAP 387
Db |||||
Qy 361 IQBFAGCKKIFGSLAFPSFDPGDPASNTAPLQ 394
Db |||||

RESULT 11

US-10-412-804A-10
; Sequence 10, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Jing, Shuguan
; APPLICANT: Ratarewicz, Suzanna
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 01-1624-A
; CURRENT APPLICATION NUMBER: US/10/412,804A
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,912
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-412-804A-10

Query Match 83.4%; Score 1873.5; DB 15; Length 715;

Best Local Similarity 89.3%; Pred. No. 2.2e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

Qy 1 MELAALCRWGLLLALLPPGAASCTVCTGDMKLRPLASPETHLDMLRHLYQGCVVQGNL 60
Db |||||
Qy 1 MELAALCRWGLLLALLPPGAASCTVCTGDMKLRPLASPETHLDMLRHLYQGCVVQGNL 60
Db |||||
Qy 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 120
Db |||||
Qy 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 120
Db |||||
Qy 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGGVLIQNPOLCYQDTILWKDIFHKNOLA 180
Db |||||
Qy 121 DPLNNTPTVGTASPGGLRELQRLSLTEILKGGVLIQNPOLCYQDTILWKDIFHKNOLA 180
Db |||||
Qy 181 LTLDITNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCCAGGCARCKGPLEPTDCCHEQC 240
Db |||||
Qy 181 LTLDITNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCCAGGCARCKGPLEPTDCCHEQC 240
Db |||||
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db |||||
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db |||||
Qy 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 357
Db |||||
Qy 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 360
Db |||||
Qy 358 XQXPA----HPVLSFLRPSWDXVSFYSPLAP 387
Db |||||
Qy 361 IQBFAGCKKIFGSLAFPSFDPGDPASNTAPLQ 394
Db |||||

RESULT 12

US-09-854-356-6
; Sequence 6, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: SmithKline Beecham Biologicals S. A.
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976

```
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein
; OTHER INFORMATION: of ECD and PD of human HER-2/neu
US-09-854-356-6

Query Match      83.4%; Score 1873.5; DB 9; Length 919;
Best Local Similarity 89.3%; Pred. No. 3e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLRLPASPEHLDMLRLHLYQGCQVVOGNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLRLPASPEHLDMLRLHLYQGCQVVOGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNVQVPLQRLRIVRGTQLPEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNVQVPLQRLRIVRGTQLPEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
QY 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGGCARCKGPIPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGGCARCKGPIPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPLENQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAVVPVXR 357
DB 301 YNYLSTDVGSCTLVCPLENQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAVVPVXR 357
QY 358 XQXPXA---HPVLSFLRPSMDXVSFYSLLAP 387
DB 361 IQEPAGCKKIFGSLAFLEPESFGDPAASNTAPLQP 394

RESULT 14
US-09-811-123-9
; Sequence 9, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ErbB
; FILE REFERENCE: GENENT.073A2
; CURRENT APPLICATION NUMBER: US/09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-123-9

Query Match      83.4%; Score 1873.5; DB 9; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLRLPASPEHLDMLRLHLYQGCQVVOGNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLRLPASPEHLDMLRLHLYQGCQVVOGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNVQVPLQRLRIVRGTQLPEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNVQVPLQRLRIVRGTQLPEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
QY 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGGCARCKGPIPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGGCARCKGPIPTDCCHQC 240

; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein
; OTHER INFORMATION: of ECD and PD of human HER-2/neu
US-09-854-356-6

Query Match      83.4%; Score 1873.5; DB 9; Length 919;
Best Local Similarity 89.3%; Pred. No. 3e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLRLPASPEHLDMLRLHLYQGCQVVOGNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLRLPASPEHLDMLRLHLYQGCQVVOGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNVQVPLQRLRIVRGTQLPEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNVQVPLQRLRIVRGTQLPEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
QY 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGGCARCKGPIPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGGCARCKGPIPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPLENQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAVVPVXR 357
DB 301 YNYLSTDVGSCTLVCPLENQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAVVPVXR 357
QY 358 XQXPXA---HPVLSFLRPSMDXVSFYSLLAP 387
DB 361 IQEPAGCKKIFGSLAFLEPESFGDPAASNTAPLQP 394

RESULT 13
US-10-146-473-72
; Sequence 72, Application US/10146473
; Publication No. US20030108888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72
; LENGTH: 1253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-72

Query Match      83.4%; Score 1873.5; DB 14; Length 1253;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;
```

Db 181 LTLDITNRSRACHPCSPMKGSRGWGSESDCQSLTRTVACGACRCKGPLPTDCCHEQC 240
Qy 241 AAGCTGPKSHSDCLACHFNHSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
Db 241 AAGCTGPKSHSDCLACHFNHSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPPLNQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAPVPX 357
Db 301 YNYLSTDVGSCTLVCPPLNQEVTAEDGTQRCCKSKPCARVYGLGMEHLREVRVTSAN 360
Qy 358 XQXPXA---HPVLSFLRPSNDXVSFAFYSLEPLAP 387
Db 361 IQEPAGCKKIFGSLAFIPESFDGDPASNTAPLQP 394

RESULT 15

US-09-811-115-3
; Sequence 3, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Brickson, Sharon
; APPLICANT: Schwalli, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3

Query Match 83.4%; Score 1873.5; DB 9; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 33; Indels 7; Gaps 2;
Qy 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPEHLDMLRHLHYQGCVVQGNL 60
Db 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPEHLDMLRHLHYQGCVVQGNL 60
Qy 61 ELTYLPTNASLSFLQDIQEVGYVLIHNOVQVPLQRLRIVRGTLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIQEVGYVLIHNOVQVPLQRLRIVRGTLFEDNYALAVLDNG 120
Qy 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
Db 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
Qy 181 LTLDITNRSRACHPCSPMKGSRGWGSESDCQSLTRTVACGACRCKGPLPTDCCHEQC 240
Db 181 LTLDITNRSRACHPCSPMKGSRGWGSESDCQSLTRTVACGACRCKGPLPTDCCHEQC 240
Qy 241 AAGCTGPKSHSDCLACHFNHSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
Db 241 AAGCTGPKSHSDCLACHFNHSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPPLNQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAPVPX 357
Db 301 YNYLSTDVGSCTLVCPPLNQEVTAEDGTQRCCKSKPCARVYGLGMEHLREVRVTSAN 360
Qy 358 XQXPXA---HPVLSFLRPSNDXVSFAFYSLEPLAP 387
Db 361 IQEPAGCKKIFGSLAFIPESFDGDPASNTAPLQP 394

RESULT 16

US-09-769-508-2
; Sequence 2, Application US/09769508

; Patent No. US20020155527A1
; GENERAL INFORMATION:
; APPLICANT: STUART, SUSAN G.
; APPLICANT: MONAHAN, JOHN J.
; APPLICANT: LANGTON, BEATRICE CLAUDIA
; APPLICANT: HANCOCK, MIRIAM E.C.
; APPLICANT: CHAO, LORRINE A.
; APPLICANT: BLUFORD, PETER
; TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75
; FILE REFERENCE: BBIO-111-C1
; CURRENT APPLICATION NUMBER: US/09/769,508
; CURRENT FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-508-2

Query Match 83.4%; Score 1873.5; DB 9; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;
Qy 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPEHLDMLRHLHYQGCVVQGNL 60
Db 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPEHLDMLRHLHYQGCVVQGNL 60
Qy 61 ELTYLPTNASLSFLQDIQEVGYVLIHNOVQVPLQRLRIVRGTLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIQEVGYVLIHNOVQVPLQRLRIVRGTLFEDNYALAVLDNG 120
Qy 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
Db 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
Qy 181 LTLDITNRSRACHPCSPMKGSRGWGSESDCQSLTRTVACGACRCKGPLPTDCCHEQC 240
Db 181 LTLDITNRSRACHPCSPMKGSRGWGSESDCQSLTRTVACGACRCKGPLPTDCCHEQC 240
Qy 241 AAGCTGPKSHSDCLACHFNHSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
Db 241 AAGCTGPKSHSDCLACHFNHSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPPLNQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAPVPX 357
Db 301 YNYLSTDVGSCTLVCPPLNQEVTAEDGTQRCCKSKPCARVYGLGMEHLREVRVTSAN 360
Qy 358 XQXPXA---HPVLSFLRPSNDXVSFAFYSLEPLAP 387
Db 361 IQEPAGCKKIFGSLAFIPESFDGDPASNTAPLQP 394

RESULT 17

US-09-854-356-1
; Sequence 1, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1

```
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: DOMAIN
; LOCATION: (1)..(653)
; OTHER INFORMATION: extracellular domain (ECD)
; NAME/KEY: DOMAIN
; LOCATION: (676)..(1255)
; OTHER INFORMATION: intracellular domain (ICD)
; NAME/KEY: DOMAIN
; LOCATION: (990)..(1255)
; OTHER INFORMATION: phosphorylation domain (PD)
; NAME/KEY: DOMAIN
; LOCATION: (990)..(1048)
; OTHER INFORMATION: fragment of the phosphorylation domain, preferred
; OTHER INFORMATION: portion (delta PD)
US-09-854-356-1

Query Match      83.4%; Score 1873.5; DB 9; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQGCQVVGNNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQGCQVVGNNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNRQVPLQRLRIVRGTQOLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNRQVPLQRLRIVRGTQOLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNRQVPLQRLRIVRGTQOLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNRQVPLQRLRIVRGTQOLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMKSGSRGSESDCQSLTRTVCAAGCCARCKGPLPTDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMKSGSRGSESDCQSLTRTVCAAGCCARCKGPLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNRQVTAEDGTQRCCKSKPCAR---GXHSXXPRAAVFPVPR 357
DB 301 YNYLSTDVGSCTLVCPHNRQVTAEDGTQRCCKSKPCAR---GXHSXXPRAAVFPVPR 357
QY 358 XQXPXA-----HPVLSFLRPSWDXVSFAFYSPLAP 387
DB 361 IQEFACKKIFGSLAFLPESFDGDPASNTAPLOP 394

RESULT 18
US-09-930-125-2
; Sequence 2, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2

QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQGCQVVGNNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQGCQVVGNNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNRQVPLQRLRIVRGTQOLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNRQVPLQRLRIVRGTQOLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMKSGSRGSESDCQSLTRTVCAAGCCARCKGPLPTDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMKSGSRGSESDCQSLTRTVCAAGCCARCKGPLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNRQVTAEDGTQRCCKSKPCAR---GXHSXXPRAAVFPVPR 357
DB 301 YNYLSTDVGSCTLVCPHNRQVTAEDGTQRCCKSKPCAR---GXHSXXPRAAVFPVPR 357
QY 358 XQXPXA-----HPVLSFLRPSWDXVSFAFYSPLAP 387
DB 361 IQEFACKKIFGSLAFLPESFDGDPASNTAPLOP 394
```

```
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-930-125-2

Query Match      83.4%; Score 1873.5; DB 9; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQGCQVVGNNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQGCQVVGNNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNRQVPLQRLRIVRGTQOLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNRQVPLQRLRIVRGTQOLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMKSGSRGSESDCQSLTRTVCAAGCCARCKGPLPTDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMKSGSRGSESDCQSLTRTVCAAGCCARCKGPLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNRQVTAEDGTQRCCKSKPCAR---GXHSXXPRAAVFPVPR 357
DB 301 YNYLSTDVGSCTLVCPHNRQVTAEDGTQRCCKSKPCAR---GXHSXXPRAAVFPVPR 357
QY 358 XQXPXA-----HPVLSFLRPSWDXVSFAFYSPLAP 387
DB 361 IQEFACKKIFGSLAFLPESFDGDPASNTAPLOP 394

RESULT 19
US-09-441-411-6
; Sequence 6, Application US/09441411
; Publication No. US20030008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Karl Erik
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441,411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-411-6

Query Match      83.4%; Score 1873.5; DB 10; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQGCQVVGNNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQGCQVVGNNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNRQVPLQRLRIVRGTQOLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNRQVPLQRLRIVRGTQOLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
```

Db 121 DPLNNTPTVTGASPGGLREQLRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQLA 180
Qy 181 LTLIDTNRSRACHPCSPMCKGSRCWGSSSDCQSLTRTVAGGCARCKGFLPTDCCHQC 240
Db 181 LTLIDTNRSRACHPCSPMCKGSRCWGSSSDCQSLTRTVAGGCARCKGFLPTDCCHQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAPVVPXR 357
Db 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
Qy 358 XQXPXA---HPVLSFLRPSWDXYSAFYSLPLAP 387
Db 361 IQEFAGCKKIFGSLAFLPESFDGDPASNTAPLQP 394

RESULT 20
US-09-984-092-4
; Sequence 4, Application US/09984092
; Publication No. US20040037840A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL THERAPEUTIC VACCINE FORMULATIONS
; FILE REFERENCE: P1011PC00
; CURRENT APPLICATION NUMBER: US/09/984,092
; CURRENT FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-092-4

Query Match 83.4%; Score 1873.5; DB 12; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

Qy 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMLRHLYQGCVVQGNL 60
Db 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMLRHLYQGCVVQGNL 60
Qy 61 ELTYLPTNASLSFLQDIQEVQGYVLIQHNQVRQVPLQRLIRVGTQLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIQEVQGYVLIQHNQVRQVPLQRLIRVGTQLFEDNYALAVLDNG 120
Qy 121 DPLNNTPTVTGASPGGLREQLRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQLA 180
Db 121 DPLNNTPTVTGASPGGLREQLRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQLA 180
Qy 181 LTLIDTNRSRACHPCSPMCKGSRCWGSSSDCQSLTRTVAGGCARCKGFLPTDCCHQC 240
Db 181 LTLIDTNRSRACHPCSPMCKGSRCWGSSSDCQSLTRTVAGGCARCKGFLPTDCCHQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAPVVPXR 357
Db 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
Qy 358 XQXPXA---HPVLSFLRPSWDXYSAFYSLPLAP 387
Db 361 IQEFAGCKKIFGSLAFLPESFDGDPASNTAPLQP 394

Publication No. US20040052811A1
; GENERAL INFORMATION:
; APPLICANT: Zielinski, Christoph
; APPLICANT: Pehamberger, Hubert
; APPLICANT: Breiteneder, Reimo
; APPLICANT: Jensen-Jarolim, Erika
; APPLICANT: Scheiner, Otto
; TITLE OF INVENTION: Vaccines Against Cancerous Diseases Associated With the HER-2/
; TITLE OF INVENTION: oncogene
; FILE REFERENCE: K 38 132/3yv
; CURRENT APPLICATION NUMBER: US/10/469,162
; CURRENT FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: PCT/EP02/02111
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: EP 01104943.4
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(675)
; OTHER INFORMATION: Extracellular Domain
US-10-469-162-3

Query Match 83.4%; Score 1873.5; DB 12; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

Qy 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMLRHLYQGCVVQGNL 60
Db 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMLRHLYQGCVVQGNL 60
Qy 61 ELTYLPTNASLSFLQDIQEVQGYVLIQHNQVRQVPLQRLIRVGTQLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIQEVQGYVLIQHNQVRQVPLQRLIRVGTQLFEDNYALAVLDNG 120
Qy 121 DPLNNTPTVTGASPGGLREQLRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQLA 180
Db 121 DPLNNTPTVTGASPGGLREQLRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQLA 180
Qy 181 LTLIDTNRSRACHPCSPMCKGSRCWGSSSDCQSLTRTVAGGCARCKGFLPTDCCHQC 240
Db 181 LTLIDTNRSRACHPCSPMCKGSRCWGSSSDCQSLTRTVAGGCARCKGFLPTDCCHQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAPVVPXR 357
Db 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
Qy 358 XQXPXA---HPVLSFLRPSWDXYSAFYSLPLAP 387
Db 361 IQEFAGCKKIFGSLAFLPESFDGDPASNTAPLQP 394

RESULT 22
US-10-253-286-553
; Sequence 553, Application US/10253286
; Publication No. US20040058881A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: II-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: 10/197,000

PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: 09/396,813
PRIOR FILING DATE: 1999-09-14
NUMBER OF SEQ ID NOS: 905
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 553
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
US-10-253-286-553

Query Match 83.4%; Score 1873.5; DB 12; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTGDMKRLPASPETHDMLRLHYQCCVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTGDMKRLPASPETHDMLRLHYQCCVQGNL 60

QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNOVQVPLQRLIRVRGTLFEDNYALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNOVQVPLQRLIRVRGTLFEDNYALAVLNG 120

QY 121 DPLNNTPTVTCASPGGLREQLRLSLTEILKGGVLIQRPOLCIQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTPTVTCASPGGLREQLRLSLTEILKGGVLIQRPOLCIQDTILWKDIFHKNQOLA 180

QY 181 LTLIDTNRSRACHPCSPMKSGRCWGSSEDCQSLTRTVCAAGGCARCKGPLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMKSGRCWGSSEDCQSLTRTVCAAGGCARCKGPLPTDCCHQC 240

QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTLVCPHNOEVTAEQGTQRCCKSKPCAR---GXHSXXPPAAVPEVXR 357
DB 301 YNYLSTDVGSCTLVCPHNOEVTAEQGTQRCCKSKPCAR---GXHSXXPPAAVPEVXR 357

QY 358 XQXPAA---HPVLSFLRPSWDVXVAFYSLPLAP 387
DB 358 XQXPAA---HPVLSFLRPSWDVXVAFYSLPLAP 387

QY 361 IQEPAGCKKIFGSLAFLPESFDGDPASNTAPLOP 394
DB 361 IQEPAGCKKIFGSLAFLPESFDGDPASNTAPLOP 394

RESULT 23
US-09-765-973-2
Sequence 2, Application US/09765973
Publication No. US2002039573A1
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Hand-Zimmermann, Susan
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR PREVENTION AND
TREATMENT OF HER-2/neu ASSOCIATED MALIGNANCIES
FILE REFERENCE: 210121.496
CURRENT APPLICATION NUMBER: US/09/765,973
CURRENT FILING DATE: 2001-01-19
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapien
US-09-765-973-2

Query Match 83.4%; Score 1873.5; DB 12; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTGDMKRLPASPETHDMLRLHYQCCVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTGDMKRLPASPETHDMLRLHYQCCVQGNL 60

QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNOVQVPLQRLIRVRGTLFEDNYALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNOVQVPLQRLIRVRGTLFEDNYALAVLNG 120

DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNOVQVPLQRLIRVRGTLFEDNYALAVLNG 120
QY 121 DPLNNTPTVTCASPGGLREQLRLSLTEILKGGVLIQRPOLCIQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTPTVTCASPGGLREQLRLSLTEILKGGVLIQRPOLCIQDTILWKDIFHKNQOLA 180

QY 181 LTLIDTNRSRACHPCSPMKSGRCWGSSEDCQSLTRTVCAAGGCARCKGPLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMKSGRCWGSSEDCQSLTRTVCAAGGCARCKGPLPTDCCHQC 240

QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTLVCPHNOEVTAEQGTQRCCKSKPCAR---GXHSXXPPAAVPEVXR 357
DB 301 YNYLSTDVGSCTLVCPHNOEVTAEQGTQRCCKSKPCAR---GXHSXXPPAAVPEVXR 357

QY 358 XQXPAA---HPVLSFLRPSWDVXVAFYSLPLAP 387
DB 358 XQXPAA---HPVLSFLRPSWDVXVAFYSLPLAP 387

QY 361 IQEPAGCKKIFGSLAFLPESFDGDPASNTAPLOP 394
DB 361 IQEPAGCKKIFGSLAFLPESFDGDPASNTAPLOP 394

RESULT 24
US-10-418-027-3
Sequence 3, Application US/10418027
Publication No. US20030224467A1
GENERAL INFORMATION:
APPLICANT: Osborne, C. Kent
APPLICANT: Schiff, Rachel
APPLICANT: Bardou, Valerie
APPLICANT: Hilsebeck, Susan
APPLICANT: Clark, Gary
APPLICANT: Wong, Jiemin
APPLICANT: Channess, Gary
APPLICANT: Hopp, Torsten
TITLE OF INVENTION: ATB 1 as a prognostic marker and predictor of endocrine therapy
TITLE OF INVENTION: resistance
FILE REFERENCE: WO-P02396US1
CURRENT APPLICATION NUMBER: US/10/418,027
CURRENT FILING DATE: 2003-04-17
PRIOR APPLICATION NUMBER: US 60/373,237
PRIOR FILING DATE: 2002-04-17
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin version 3.1
SEQ ID NO 3
LENGTH: 1255
TYPE: PRT
ORGANISM: Human
US-10-418-027-3

Query Match 83.4%; Score 1873.5; DB 12; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTGDMKRLPASPETHDMLRLHYQCCVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTGDMKRLPASPETHDMLRLHYQCCVQGNL 60

QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNOVQVPLQRLIRVRGTLFEDNYALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNOVQVPLQRLIRVRGTLFEDNYALAVLNG 120

QY 121 DPLNNTPTVTCASPGGLREQLRLSLTEILKGGVLIQRPOLCIQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTPTVTCASPGGLREQLRLSLTEILKGGVLIQRPOLCIQDTILWKDIFHKNQOLA 180

QY 181 LTLIDTNRSRACHPCSPMKSGRCWGSSEDCQSLTRTVCAAGGCARCKGPLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMKSGRCWGSSEDCQSLTRTVCAAGGCARCKGPLPTDCCHQC 240

QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHNOBVTAEADGTQRCCKSKPCAR---GXHSXXPRPAAPVPVXR 357
Db 301 YNYLSTDVGSCTLVCPHNOBVTAEADGTQRCCKSKPCARVCYGLGMEHLREVRAVTSAN 360
Qy 358 XQXPA---HPVLSFLRPSWDXVSAYFSLPLAP 387
Db 361 IQEPAGCKKIFGSLAFLPESFDGDPASNTAPLQP 394

RESULT 25

US-10-207-655-45
; Sequence 45, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390693.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 45
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-45

Query Match 83.4%; Score 1873.5; DB 14; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

Qy 1 MELAALCRWGLLALLPPGAASQVCTGTDMKRLPASPETHLDMRLHLYQGCVVQGNL 60
Db 1 MELAALCRWGLLALLPPGAASQVCTGTDMKRLPASPETHLDMRLHLYQGCVVQGNL 60
Qy 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNOBVTAEADGTQRCCKSKPCAR---GXHSXXPRPAAPVPVXR 120
Db 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNOBVTAEADGTQRCCKSKPCAR---GXHSXXPRPAAPVPVXR 120
Qy 121 DPLNNTTPTVGTASPGGLREQLRLSLEILKGGVLIQRPOLCYQDTILWKDI FHKNNOLA 180
Db 121 DPLNNTTPTVGTASPGGLREQLRLSLEILKGGVLIQRPOLCYQDTILWKDI FHKNNOLA 180
Qy 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCGAGCARCKGPLPTDCCHEQC 240
Db 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCGAGCARCKGPLPTDCCHEQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHNOBVTAEADGTQRCCKSKPCAR---GXHSXXPRPAAPVPVXR 357
Db 301 YNYLSTDVGSCTLVCPHNOBVTAEADGTQRCCKSKPCARVCYGLGMEHLREVRAVTSAN 360
Qy 358 XQXPA---HPVLSFLRPSWDXVSAYFSLPLAP 387
Db 361 IQEPAGCKKIFGSLAFLPESFDGDPASNTAPLQP 394

RESULT 26

US-10-177-293-126
; Sequence 126, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lallie, James
; APPLICANT: Giatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula

; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel B.
; APPLICANT: East Jr., Robert C.
; APPLICANT: Hortobagyi, Gabrieli N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-036
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-126

Query Match 83.4%; Score 1873.5; DB 14; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

Qy 1 MELAALCRWGLLALLPPGAASQVCTGTDMKRLPASPETHLDMRLHLYQGCVVQGNL 60
Db 1 MELAALCRWGLLALLPPGAASQVCTGTDMKRLPASPETHLDMRLHLYQGCVVQGNL 60
Qy 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNOBVTAEADGTQRCCKSKPCAR---GXHSXXPRPAAPVPVXR 120
Db 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNOBVTAEADGTQRCCKSKPCAR---GXHSXXPRPAAPVPVXR 120
Qy 121 DPLNNTTPTVGTASPGGLREQLRLSLEILKGGVLIQRPOLCYQDTILWKDI FHKNNOLA 180
Db 121 DPLNNTTPTVGTASPGGLREQLRLSLEILKGGVLIQRPOLCYQDTILWKDI FHKNNOLA 180
Qy 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCGAGCARCKGPLPTDCCHEQC 240
Db 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCGAGCARCKGPLPTDCCHEQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHNOBVTAEADGTQRCCKSKPCAR---GXHSXXPRPAAPVPVXR 357
Db 301 YNYLSTDVGSCTLVCPHNOBVTAEADGTQRCCKSKPCARVCYGLGMEHLREVRAVTSAN 360
Qy 358 XQXPA---HPVLSFLRPSWDXVSAYFSLPLAP 387
Db 361 IQEPAGCKKIFGSLAFLPESFDGDPASNTAPLQP 394

RESULT 27

```
US-10-207-498-6
; Sequence 6, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HEREULIN AND HER3
; FILE REFERENCE: 30448.103-US-J1
; CURRENT APPLICATION NUMBER: US/10/207,498
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-498-6

Query Match      83.4%; Score 1873.5; DB 14; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMLRHLRYGCGQVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMLRHLRYGCGQVQGNL 60

QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIAHNOVQVPLQRLIRVRGTLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIAHNOVQVPLQRLIRVRGTLFEDNYALAVLDNG 120

QY 121 DPLANNTPVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
DB 121 DPLANNTPVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180

QY 181 LTLIDTNRSRACHPCSPCKSGRCWGESSEDCQSLTRTVGAGGCARCKGPIPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPCKSGRCWGESSEDCQSLTRTVGAGGCARCKGPIPTDCCHQC 240

QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTLVCPNLNQEVTAEDGTQRCCKSKPCAR---GXHXXXPRPAAVVPVEXR 357
DB 301 YNYLSTDVGSCTLVCPNLNQEVTAEDGTQRCCKSKPCAR---GXHXXXPRPAAVVPVEXR 357

QY 358 XQXPFA---HPVLFLRPSMDXVSFAFYSPLAP 387
DB 361 IQEPAGCKKIFGSLAFIPESFGDPASNTAPLQP 394

RESULT 29
US-10-313-644-2
; Sequence 2, Application US/10313644
; Publication No. US20030157119A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: Hand-Zimmerman, Susan
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
; TITLE OF INVENTION: AND VIRUS-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.483C3
; CURRENT APPLICATION NUMBER: US/10/313,644
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-313-644-2

Query Match      83.4%; Score 1873.5; DB 14; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-154;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMLRHLRYGCGQVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMLRHLRYGCGQVQGNL 60

QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIAHNOVQVPLQRLIRVRGTLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIAHNOVQVPLQRLIRVRGTLFEDNYALAVLDNG 120

QY 121 DPLANNTPVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
DB 121 DPLANNTPVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180

US-10-338-730-2
; Sequence 2, Application US/10338730
; Publication No. US20030147905A1
; GENERAL INFORMATION:
; APPLICANT: Genzyme Corporation
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS
; FILE REFERENCE: 5017C
; CURRENT APPLICATION NUMBER: US/10/338,730
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: US 09/527,487
; PRIOR FILING DATE: 2002-03-16
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1255
```


QY 181 LTLIDTNRSRACHPCSPMKGSRGWGSESDCQSLTRTVCAAGGCARCKGPLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMKGSRGWGSESDCQSLTRTVCAAGGCARCKGPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLAHFNHSGICELHCPALVTNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLAHFNHSGICELHCPALVTNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAVVPVXR 357
DB 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
QY 358 XQXPA---HPVLSFLRPSWDVSAFYSLPLAP 387
DB 361 IQFAGCKKIFGSLAFIPESFDGDPASNTAPLQP 394

RESULT 30
US-10-322-892-4
; Sequence 4, Application US/10322892
; Publication No. US20030171257A1
; GENERAL INFORMATION:
; APPLICANT: STIRBL, ROBERT C.
; APPLICANT: SNEAD, MALCOLM L.
; APPLICANT: XU, JIMMY
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: MILK, PETER J.
; TITLE OF INVENTION: METHOD AND RELATED COMPOSITION EMPLOYING NANOSTRUCTURES
; FILE REFERENCE: W07-505
; CURRENT APPLICATION NUMBER: US/10/322,892
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: 60/342,894
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-322-892-4

Query Match 83.4%; Score 1873.5; DB 14; Length 1255;
Best Local Similarity 89.3%; Pred. No. 4.5e-134;
Matches 352; Conservative 3; Mismatches 32; Indels 7; Gaps 2;
QY 1 MELAALCFWGLLLALLPFGAASCTQCTGTMKLELPASPETHDMLRHLVGGCQVQGNL 60
DB 1 MELAALCFWGLLLALLPFGAASCTQCTGTMKLELPASPETHDMLRHLVGGCQVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIAHNQVRQVPLQRLRIVRGTLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIAHNQVRQVPLQRLRIVRGTLFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTCASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTTPTVTCASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPMKGSRGWGSESDCQSLTRTVCAAGGCARCKGPLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMKGSRGWGSESDCQSLTRTVCAAGGCARCKGPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLAHFNHSGICELHCPALVTNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLAHFNHSGICELHCPALVTNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR---GXHSXXPRPAAVVPVXR 357
DB 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
QY 358 XQXPA---HPVLSFLRPSWDVSAFYSLPLAP 387
DB 361 IQFAGCKKIFGSLAFIPESFDGDPASNTAPLQP 394

Search completed: July 4, 2004, 04:16:55
Job time : 87.8193 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2004, 04:09:33 ; Search time 16.3394 Seconds
(without alignments)
1366.103 Million cell updates/sec

Title: US-09-506-079H-11

Perfect score: 418
Sequence: 1 GTHSLPRAAPVFLRMQP.....VGRGPDPAHVAVNLSRYEG 79

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 90 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001s.*

5: Geneseq2002s.*

6: Geneseq2003as.*

7: Geneseq2003bs.*

8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	418	100.0	79	4 AAE09196	Aae09196 Human p68
2	418	100.0	83	4 AAE09185	Aae09185 Human p68
3	418	100.0	84	5 AAE20349	Aae20349 HER2 prot
4	413	98.8	79	4 AAE09188	Aae09188 Human p68
5	411	98.3	79	4 AAE09195	Aae09195 Human p68
6	411	98.3	79	4 AAE09199	Aae09199 Human p68
7	406	97.1	79	3 AAY97241	Aay97241 HER-2 C-t
8	406	97.1	79	4 AAE09184	Aae09184 Human p68
9	406	97.1	79	5 AAE20350	Aae20350 Human HER
10	404	96.7	79	4 AAE09192	Aae09192 Human p68
11	403	96.4	79	4 AAE09190	Aae09190 Human p68
12	402	96.2	79	4 AAE09186	Aae09186 Human p68
13	400	95.7	79	4 AAE09189	Aae09189 Human p68
14	399	95.5	79	4 AAE09191	Aae09191 Human p68
15	399	95.5	79	4 AAE09187	Aae09187 Human p68
16	398	95.2	79	4 AAE09197	Aae09197 Human p68
17	398	95.2	79	4 AAE09198	Aae09198 Human p68
18	397	95.0	79	4 AAE09193	Aae09193 Human p68
19	396	94.7	79	4 AAE09194	Aae09194 Human p68
20	363	86.8	419	4 AAE09213	Aae09213 Human p68
21	358	85.6	419	4 AAE09212	Aae09212 Human p68
22	356	85.2	419	4 AAE09205	Aae09205 Human p68
23	355	84.9	79	3 AAY97239	Aay97239 HER-2 C-t
24	355	84.9	79	4 AAE09180	Aae09180 Human p68
25	355	84.9	79	5 AAE20347	Aae20347 Human HER

ALIGNMENTS

RESULT 1

AAE09196

ID AAE09196 standard; peptide; 79 AA.

26	354	84.7	419	4 AAE09207	Aae09207 Human p68
27	354	84.7	419	4 AAE09209	Aae09209 Human p68
28	352	84.2	419	4 AAE09203	Aae09203 Human p68
29	351	84.0	419	4 AAE09210	Aae09210 Human p68
30	351	84.0	419	4 AAE09208	Aae09208 Human p68
31	351	84.0	419	4 AAE09181	Aae09181 Human p68
32	351	84.0	419	5 AAE20348	Aae20348 Human tru
33	351	84.0	420	3 AAY97240	Aay97240 Truncated
34	350	83.7	419	4 AAE09211	Aae09211 Human p68
35	350	83.7	419	4 AAE09206	Aae09206 Human p68
36	349	83.5	419	4 AAE09204	Aae09204 Human p68
37	346	82.8	419	4 AAE09216	Aae09216 Human p68
38	343	82.1	79	4 AAE09182	Aae09182 Human p68
39	343	82.1	419	4 AAE09202	Aae09202 Human p68
40	343	82.1	419	4 AAE09200	Aae09200 Human p68
41	339	81.1	419	4 AAE09183	Aae09183 Human p68
42	337	80.6	419	4 AAE09214	Aae09214 Human p68
43	337	80.6	419	4 AAE09215	Aae09215 Human p68
44	86	20.6	459	2 AAY31404	Aay31404 HEV US an
45	86	20.6	459	2 AAY31405	Aay31405 HEV US re
46	83.5	20.0	888	2 AAW19857	Aaw19857 Human sem
47	83.5	20.0	888	3 AAB18916	Aab18916 A novel p
48	83.5	20.0	888	4 AAU12443	Aau12443 Human PRO
49	83.5	20.0	888	5 ABB84950	Abb84950 Human PRO
50	83.5	20.0	888	5 ABB95556	Abb95556 Human ang
51	83.5	20.0	888	6 ABO03538	Abu03538 Argigene
52	83.5	20.0	888	6 ABO17887	Abu17887 Novel hum
53	83.5	20.0	888	6 ABO69110	Abu69110 Human PRO
54	83.5	20.0	888	6 ABO81141	Abu81141 Human PRO
55	83.5	20.0	888	6 ABO19426	Abu19426 Human sec
56	83.5	20.0	888	6 ABU66841	Abu66841 Human PRO
57	83.5	20.0	888	6 ABU59922	Abu59922 Novel sec
58	83.5	20.0	888	6 ABU69087	Abu69087 Human PRO
59	83.5	20.0	888	6 ABO25112	Abu25112 Human sec
60	83.5	20.0	888	6 ABU67117	Abu67117 Human sec
61	83.5	20.0	888	6 ABU81551	Abu81551 Human sec
62	83.5	20.0	888	6 ADA46063	Ada46063 Novel hum
63	83.5	20.0	888	6 ADA76560	Ada76560 Novel hum
64	83.5	20.0	888	6 ADA76494	Ada76494 Human PRO
65	83.5	20.0	888	6 ADA19144	Ada19144 Human PRO
66	83.5	20.0	888	6 ADA61767	Ada61767 Homo sapi
67	83.5	20.0	888	6 ADB19552	Adb19552 Novel hum
68	83.5	20.0	888	6 ADB28093	Adb28093 Human PRO
69	83.5	20.0	888	6 ADA86572	Ada86572 Novel hum
70	83.5	20.0	888	6 ADB16136	Adb16136 Human PRO
71	83.5	20.0	888	6 ADA47922	Ada47922 Human PRO
72	83.5	20.0	888	6 ADA67717	Ada67717 Human PRO
73	83.5	20.0	888	6 ADB30724	Adb30724 Human PRO
74	83.5	20.0	888	6 ADA86020	Ada86020 Novel hum
75	83.5	20.0	888	6 ADA97232	Ada97232 Human PRO
76	83.5	20.0	888	6 ADA79536	Ada79536 Human PRO
77	83.5	20.0	888	6 ADA87675	Ada87675 Novel hum
78	83.5	20.0	888	6 ADB16877	Adb16877 Human PRO
79	83.5	20.0	888	6 ADA91969	Ada91969 Novel hum
80	83.5	20.0	888	6 ADB15032	Adb15032 Human PRO
81	83.5	20.0	888	6 ADB18993	Adb18993 Novel hum
82	83.5	20.0	888	6 ADA94208	Ada94208 Human PRO
83	83.5	20.0	888	6 ADB20104	Adb20104 Novel hum
84	83.5	20.0	888	6 ADB13416	Adb13416 Human PRO
85	83.5	20.0	888	6 ABO43420	Abu43420 Novel hum
86	83.5	20.0	888	6 ADA74670	Ada74670 Human PRO
87	83.5	20.0	888	6 ADB24903	Adb24903 Human PRO
88	83.5	20.0	888	6 ADA82427	Ada82427 Human PRO
89	83.5	20.0	888	6 ADA75390	Ada75390 Human PRO
90	83.5	20.0	888	6 ADA85468	Ada85468 Novel hum

```

XX AC AAE09196;
XX AC AAE09185 standard; protein; 83 AA.
XX DT 15-NOV-2001 (first entry)
XX AC AAE09185;
XX DT 15-NOV-2001 (first entry)
XX DE Human p68HER-2 ECDIIIA variant 11 encoded by HER-2 intron 8.
XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
XX KW ECDIIIA; variant.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
XX FT Misc-difference 6 /note= "p68HER-2 ECDIIIA (AAE09184) Pro substituted with
XX FT Leu"
XX FT Misc-difference 73 /note= "p68HER-2 ECDIIIA (AAE09184) Asp substituted with
XX FT Asn which is encoded by CAC"
XX PN W0200161356-A1.
XX XX 23-AUG-2001.
XX PF 16-FEB-2001; 2001WO-US005327.
XX PR 16-FEB-2000; 2000US-00506079.
XX XX (UYOR-) UNIV OREGON HEALTH SCI.
XX PI Clinton G, Henner WD, Evans A;
XX DR WPI; 2001-529934/58.
XX DR N-PSDB; AAD15865.
XX XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX FT treatment of hard tumors.
XX PS Example 11; Page; 61pp; English.
XX CC The invention relates to novel HER-2 (herstatin-2) antagonist
XX CC particularly a polypeptide that binds to the extracellular domain (ECD)
XX CC of HER-2 at a site that is different from the binding site of humanised
XX CC antibody, Herceptin, at an affinity of at least 108. The present
XX CC invention is based upon the initial discovery of an alternative HER-2
XX CC mRNA transcript with 274 bp insert of intron 8. The translation product
XX CC of the alternative transcript is a truncated HER-2 protein designated
XX CC p68HER-2 which lacks the transmembrane and intracellular domains of
XX CC p68HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX CC nucleic acids encoding these are useful to treat, diagnose and identify
XX CC solid tumours. The present sequence is p68HER-2 ECDIIIA variant encoded
XX CC by polymorphic form of human HER-2 intron 8. Note: The present sequence
XX CC is not shown in the specification but is derived from HER-2 intron 8
XX CC encoded ECDIIIA sequence given in figure 8 (AAE09184).
XX XX Sequence 79 AA;
XX SQ Query Match 100.0%; Score 418; DB 4; Length 79;
XX Best Local Similarity 100.0%; Pred. No. 2.2e-38;
XX Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTHSLPRPAAPVPLRMQPGPAHPVLSFLRPSWDLVSAYFSLPLAPLSPTSPVSPSV 60
DB 1 GTHSLPRPAAPVPLRMQPGPAHPVLSFLRPSWDLVSAYFSLPLAPLSPTSPVSPSV 60
QY 61 GRGPDPAHVAVNLRYEG 79
DB 61 GRGPDPAHVAVNLRYEG 79

```

```

RESULT 2
AAE09185
ID AAE09185 standard; protein; 83 AA.
XX AC AAE09185;
XX DT 15-NOV-2001 (first entry)
XX DE Human p68HER-2 partial protein containing ECDIIIA variant sequence;
XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
XX KW ECDIIIA; variant.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
XX FT Misc-difference 1 /note= "Encoded by CC"
XX FT Domain 3.81 /label= ECDIIIA
XX FT /note= "Extracellular domain IIA of p68HER-2"
XX FT Misc-difference 8 /note= "Most commonly occurring ECDIIIA (AAE09184) Pro
XX FT substituted with Leu"
XX FT Modified-site 75
XX FT /note= "Asn is N-glycosylated. Most commonly occurring
XX FT ECDIIIA (AAE09184) Asp substituted with Asn"
XX PN W0200161356-A1.
XX XX 23-AUG-2001.
XX PF 16-FEB-2001; 2001WO-US005327.
XX PR 16-FEB-2000; 2000US-00506079.
XX XX (UYOR-) UNIV OREGON HEALTH SCI.
XX PI Clinton G, Henner WD, Evans A;
XX DR WPI; 2001-529934/58.
XX DR N-PSDB; AAD15854.
XX XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX FT treatment of hard tumors.
XX PS Example 1; Fig 1; 61pp; English.
XX CC The invention relates to novel HER-2 (herstatin-2) antagonist
XX CC particularly a polypeptide that binds to the extracellular domain (ECD)
XX CC of HER-2 at a site that is different from the binding site of humanised
XX CC antibody, Herceptin, at an affinity of at least 108. The present
XX CC invention is based upon the initial discovery of an alternative HER-2
XX CC mRNA transcript with 274 bp insert of intron 8. The translation product
XX CC of the alternative transcript is a truncated HER-2 protein designated
XX CC p68HER-2 which lacks the transmembrane and intracellular domains of
XX CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX CC nucleic acids encoding these are useful to treat, diagnose and identify
XX CC solid tumours. The present sequence is human p68HER-2 partial protein
XX CC containing ECDIIIA variant sequence
XX XX Sequence 83 AA;
XX SQ Query Match 100.0%; Score 418; DB 4; Length 83;
XX Best Local Similarity 100.0%; Pred. No. 2.3e-38;
XX Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTHSLPRPAAPVPLRMQPGPAHPVLSFLRPSWDLVSAYFSLPLAPLSPTSPVSPSV 60
DB 3 GTHSLPRPAAPVPLRMQPGPAHPVLSFLRPSWDLVSAYFSLPLAPLSPTSPVSPSV 62

```


QY 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVSPSV 60
DB 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVSPSV 60
QY 61 GRGPDPAHVAVNLSRYEG 79
DB 61 GRGPDPAHVAVNLSRYEG 79

RESULT 5
AAE09195
ID AAE09195 standard; peptide; 79 AA.

XX AC AAE09195;
XX DT 15-NOV-2001 (first entry)
XX DE Human p68HER-2 ECDIIa variant 10 encoded by HER-2 intron 8.
XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIa; variant.
XX OS Homo sapiens.

XX FH Key Location/Qualifiers
XX FT Misc-difference /note= "p68HER-2 ECDIIa (AAE09184) Asp substituted with
FT Asn which is encoded by CAC"
XX PN WO200161356-A1.

XX PD 23-AUG-2001.
XX PF 16-FEB-2001; 2001WO-US005327.
XX PR 16-FEB-2000; 2000US-00506079.
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX PI Clinton G, Henner WD, Evans A;
XX WPI; 2001-529334/58.
XX DR N-PSDB; AAD15864.

XX PT New polypeptide, which binds to the extracellular domain of HER-2 for the
PT treatment of hard tumors.
XX PS Example 11; Page; 61pp; English.

XX CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10⁸. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p68HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDIIa.
CC The ECDIIa-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is p68HER-2 ECDIIa variant encoded
CC by polymorphic form of human HER-2 intron 8. Note: The present sequence
CC is not shown in the specification but is derived from HER-2 intron 8
CC encoded ECDIIa sequence given in figure 8 (AAE09184)

XX SQ Sequence 79 AA;
Query Match 98.3%; Score 411; DB 4; Length 79;
Best Local Similarity 98.7%; Pred. No. 1.3e-37;
Matches 78; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVSPSV 60
DB 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVSPSV 60
QY 61 GRGPDPAHVAVNLSRYEG 79
DB 61 GRGPDPAHVAVNLSRYEG 79

RESULT 6
AAE09199
ID AAE09199 standard; peptide; 79 AA.

XX AC AAE09199;
XX DT 15-NOV-2001 (first entry)
XX DE Human p68HER-2 ECDIIa variant 15 encoded by HER-2 intron 8.
XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIa; variant.
XX OS Homo sapiens.

XX FH Key Location/Qualifiers
XX FT Misc-difference /note= "p68HER-2 ECDIIa (AAE09184) Asp substituted with
FT Asn"
XX PN WO200161356-A1.

XX PD 23-AUG-2001.
XX PF 16-FEB-2001; 2001WO-US005327.
XX PR 16-FEB-2000; 2000US-00506079.
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX PI Clinton G, Henner WD, Evans A;
XX WPI; 2001-529334/58.
XX DR N-PSDB; AAD15868.
XX PT New polypeptide, which binds to the extracellular domain of HER-2 for the
PT treatment of hard tumors.
XX PS Example 12; Page; 61pp; English.

XX CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10⁸. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p68HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDIIa.
CC The ECDIIa-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is p68HER-2 ECDIIa variant encoded
CC by polymorphic form of human HER-2 intron 8. Note: The present sequence
CC is not shown in the specification but is derived from HER-2 intron 8
CC encoded ECDIIa sequence given in figure 8 (AAE09184)

XX SQ Sequence 79 AA;
Query Match 98.3%; Score 411; DB 4; Length 79;
Best Local Similarity 98.7%; Pred. No. 1.3e-37;
Matches 78; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVISPVS 60
Db 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVISPVS 60
QY 61 GRGPDPAHVAVNLSRYEG 79
Db 61 GRGPDPAHVAVNLSRYEG 79

RESULT 7
AA97241
ID AA97241 standard; protein; 79 AA.
XX AC AA97241;
XX DT 04-DEC-2000 (first entry)
XX DE HER-2 C-terminal extracellular domain IIIa.
XX KW HER-2; erbB-2; oncogene; receptor-like tyrosine kinase; insertion;
XX KW extracellular domain IIIa; antagonist; intron 8; C-terminal extension;
XX KW truncated HER-2; p68; dimerization inhibitor; cytostatic.
XX OS Homo sapiens.
XX PN WO200044403-A1.
XX PD 03-AUG-2000.
XX PF 20-JAN-2000; 2000WO-US001484.
XX PR 20-JAN-1999; 95US-00234208.
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX PI Doherty JK, Clinton GM, Adelman JP;
XX WPI; 2000-499287/44.
XX DR N-PSDB; AA53783.
XX PT Using polypeptides and antibodies that bind to the extracellular domain
XX PT of the receptor-like tyrosine kinase HER-2 to treat solid tumors of the
XX PT breast, lung, ovaries and colon.
XX PS Example 11; Page 42-43; 45pp; English.
XX CC HER-2/neu (erbB-2) oncogene encodes a receptor-like tyrosine kinase. The
XX CC extracellular domain of p185-HER-2 is proteolytically shed from breast
XX CC carcinoma cells in culture and is found in serum of some cancer patients
XX CC and may be a serum marker of metastatic breast cancer. An alternative HER
XX CC -2 mRNA of 4.8 kb with a 274 bp insert (intron 8) has been identified.
XX CC The retained intron is in-frame and encodes a 79 amino acid extension.
XX CC designated ECDIIIA (the present sequence), which is inserted at residue
XX CC 340 of p185-HER-2. The alternative mRNA predicts a truncated HER-2
XX CC protein (approximately 68 kDa) that lacks the transmembrane and
XX CC intracellular domains (see AA97240). p68HER-2 specifically binds to p185
XX CC -HER-2 without activating HER-2. It could therefore block dimerization of
XX CC p185-HER-2. The p68HER-2 polypeptide binds to a site on the ECD of HER-2
XX CC that is different from the site of binding for Herceptin (RPM) (a
XX CC marketed humanized monoclonal antibody that is used for the treatment of
XX CC cancer and binds to the ECD of HER-2). The methods, compositions,
XX CC polypeptides and antibodies are used to treat solid tumors such as
XX CC breast cancer, small cell lung carcinoma, ovarian cancer and/or colon
XX CC cancer, especially where over-expression of HER-2 is indicated
XX SQ Sequence 79 AA;

Query Match 97.1%; Score 406; DB 3; Length 79;
Best Local Similarity 97.5%; Pred. No. 4.5e-37;
Matches 77; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVISPVS 60
Db 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVISPVS 60
```

```
Db 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVISPVS 60
QY 61 GRGPDPAHVAVNLSRYEG 79
Db 61 GRGPDPAHVAVNLSRYEG 79

RESULT 8
AAE09184
ID AAE09184 standard; peptide; 79 AA.
XX AC AAE09184;
XX DT 15-NOV-2001 (first entry)
XX DE Human p68HER-2 ECDIIIA domain encoded by HER-2 intron 8.
XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
XX KW ECDIIIA; intron 8.
XX OS Homo sapiens.
XX PN WO200161356-A1.
XX PD 23-AUG-2001.
XX PF 16-FEB-2001; 2001WO-US005327.
XX PR 16-FEB-2000; 2000US-00506079.
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX PI Clinton G, Henner WD, Evans A;
XX WPI; 2001-529934/58.
XX DR N-PSDB; AAD15853, AAD15869.
XX PT New polypeptide, which binds to the extracellular domain of HER-2 for the
XX PT treatment of hard tumors.
XX PS Example 11; Fig 8; 61pp; English.
XX CC The invention relates to novel HER-2 (herstatin-2) antagonist
XX CC particularly a polypeptide that binds to the extracellular domain (ECD)
XX CC of HER-2 at a site that is different from the binding site of humanised
XX CC antibody, Herceptin, at an affinity of at least 108. The present
XX CC invention is based upon the initial discovery of an alternative HER-2
XX CC mRNA transcript with 274 bp insert of intron 8. The translation product
XX CC of the alternative transcript is a truncated HER-2 protein designated
XX CC p68HER-2 which lacks the transmembrane and intracellular domains of
XX CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX CC the HER-2 receptor, the peptides, which bind to an HER-2 ECD, and the
XX CC nucleic acids encoding these are useful to treat, diagnose and identify
XX CC solid tumours. The present sequence is ECDIIIA domain of human p68HER-2
XX CC protein encoded by intron 8 of HER-2 gene
XX SQ Sequence 79 AA;

Query Match 97.1%; Score 406; DB 4; Length 79;
Best Local Similarity 97.5%; Pred. No. 4.5e-37;
Matches 77; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVISPVS 60
Db 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVISPVS 60
QY 61 GRGPDPAHVAVNLSRYEG 79
Db 61 GRGPDPAHVAVNLSRYEG 79
```

```
RESULT 9
AAE020350
ID AAE020350 standard; protein; 79 AA.
XX
AC AAE020350;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human HER2 intron 8 encoded protein.
XX
KW Human; tumour; endothelial growth factor receptor; EGFR; cytostatic;
KW herstatin; HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung;
KW colon; glial cell tumour; cell growth.
XX
OS Homo sapiens.
XX
PN WO200214470-A2.
XX
PD 21-FEB-2002.
XX
PF 14-AUG-2001; 2001WO-US025502.
XX
PR 14-AUG-2000; 2000US-00638834.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Clinton GM;
XX
DR WPI; 2002-269185/31.
XX
DR N-PSDB; AAD32540.
XX
PT Treating solid tumor characterized by expression of endothelial growth
PT factor receptor, involves administering recombinant herstatin that binds
PT to extracellular domain of the endothelial growth factor receptor.
XX
PS Example 11; Fig 8; 82pp; English.
XX
CC The present invention relates to a method for treating a solid tumour
CC characterised by endothelial growth factor receptor (EGFR) expression.
CC The method involves administering an agent that binds to an extracellular
CC domain (ECD) of EGFR. The invention also relates to a naturally occurring
CC inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-
CC expression of herstatin with p185HER2 causes a striking reduction in cell
CC growth that corresponds with suppression of p185 autophosphorylation. The
CC method or a pharmaceutical composition is useful for treating a solid
CC tumour (selected from squamous cell carcinoma, lung carcinoma, colon
CC carcinoma and glial cell tumour) characterised by EGFR expression. The
CC present sequence is a protein encoded by human HER2 intron 8
XX
SQ Sequence 79 AA;
Query Match 97.1%; Score 406; DB 5; Length 79;
Best Local Similarity 97.5%; Pred. No. 4.5e-37;
Matches 77; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GTHSLLPRAAVPVPLRMQPGPAHVLFLRPSMDLVSAFYSLPLAPLSPTSVPISPSV 60
DB 1 GTHSLFPRAAVPVPLRMQPGPAHVLFLRPSMDLVSAFYSLPLAPLSPTSVPISPSV 60
QY 61 GRGPDPAHVAVNLKRYEG 79
DB 61 GRGPDPAHVAVDLSRYEG 79
RESULT 10
AAE09192
ID AAE09192 standard; peptide; 79 AA.
XX
AC AAE09192;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 ECDIIIA variant 7 encoded by HER-2 intron 8.
```

```
XX
KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 36 /note= "p68HER-2 ECDIIIA (AAE09184) Leu substituted with
FT FT Ile"
PN WO200161356-A1.
XX
PD 23-AUG-2001.
XX
PF 16-FEB-2001; 2001WO-US005327.
XX
PR 16-FEB-2000; 2000US-00506079.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Clinton G, Henner WD, Evans A;
XX
DR WPI; 2001-529934/58.
XX
DR N-PSDB; AAD15861.
XX
PT New polypeptide, which binds to the extracellular domain of HER-2 for the
PT treatment of hard tumors.
XX
PS Example 11; Page; 61pp; English.
XX
CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 108. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC cDNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumors. The present sequence is p68HER-2 ECDIIIA variant encoded
CC by polymorphic form of human HER-2 intron 8. Note: The present sequence
CC is not shown in the specification but is derived from HER-2 intron 8
CC encoded ECDIIIA sequence given in figure 8 (AAE09184)
XX
SQ Sequence 79 AA;
Query Match 96.7%; Score 404; DB 4; Length 79;
Best Local Similarity 96.2%; Pred. No. 7.5e-37;
Matches 76; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 GTHSLLPRAAVPVPLRMQPGPAHVLFLRPSMDLVSAFYSLPLAPLSPTSVPISPSV 60
DB 1 GTHSLFPRAAVPVPLRMQPGPAHVLFLRPSMDLVSAFYSLPLAPLSPTSVPISPSV 60
QY 61 GRGPDPAHVAVNLKRYEG 79
DB 61 GRGPDPAHVAVDLSRYEG 79
RESULT 11
AAE09190
ID AAE09190 standard; peptide; 79 AA.
XX
AC AAE09190;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 ECDIIIA variant 5 encoded by HER-2 intron 8.
```

```

XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 18
XX FT /note= "p68HER-2 ECDIIIA (AAE09184) Met substituted with
XX FT Leu which is encoded by ATA"
XX
XX PN WO200161356-A1.
XX
XX PD 23-AUG-2001.
XX
XX PF 16-FEB-2001; 2001WO-US0053327.
XX
XX PR 16-FEB-2000; 2000US-00506079.
XX
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX PI Clinton G, Henner WD, Evans A;
XX
XX DR WPI; 2001-529934/58.
XX DR N-PSDB; AAD15859.
XX
XX PT New polypeptide, which binds to the extracellular domain of HER-2 for the
XX PT treatment of hard tumors.
XX
XX PS Example 11; Page; 61pp; English.
XX
XX CC The invention relates to novel HER-2 (herstatin-2) antagonist
XX CC particularly a polypeptide that binds to the extracellular domain (ECD)
XX CC of HER-2 at a site that is different from the binding site of humanised
XX CC antibody, Herceptin, at an affinity of at least 108. The present
XX CC invention is based upon the initial discovery of an alternative HER-2
XX CC mRNA transcript with 274 bp insert of intron 8. The translation product
XX CC of the alternative transcript is a truncated HER-2 protein designated
XX CC p68HER-2 which lacks the transmembrane and intracellular domains of
XX CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX CC nucleic acids encoding these are useful to treat, diagnose and identify
XX CC solid tumours. The present sequence is p68HER-2 ECDIIIA variant encoded
XX CC by polymorphic form of human HER-2 intron 8. Note: The present sequence
XX CC is not shown in the specification but is derived from HER-2 intron 8
XX CC encoded ECDIIIA sequence given in figure 8 (AAE09184)
XX
XX SQ Sequence 79 AA;
XX
XX Query Match 96.4%; Score 403; DB 4; Length 79;
XX Best Local Similarity 96.2%; Pred. No. 9.6e-37;
XX Matches 76; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 GTHSLPPRPAAPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVSPVSV 60
XX Db 1 GTHSLPPRPAAPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVSPVSV 60
XX
XX QY 61 GRGPDPAHVAVNLSRYEG 79
XX Db 61 GRGPDPAHVAVNLSRYEG 79
XX
XX RESULT 12
XX AAE09186
XX ID AAE09186 standard; peptide; 79 AA.
XX
XX AC AAE09186;
XX
XX DT 15-NOV-2001 (first entry)
XX
XX DE Human p68HER-2 ECDIIIA variant 1 encoded by HER-2 intron 8.

```

```

XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 2
XX FT /note= "p68HER-2 ECDIIIA (AAE09184) Thr substituted with
XX FT Ser"
XX
XX PN WO200161356-A1.
XX
XX PD 23-AUG-2001.
XX
XX PF 16-FEB-2001; 2001WO-US0053327.
XX
XX PR 16-FEB-2000; 2000US-00506079.
XX
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX PI Clinton G, Henner WD, Evans A;
XX
XX DR WPI; 2001-529934/58.
XX DR N-PSDB; AAD15855.
XX
XX PT New polypeptide, which binds to the extracellular domain of HER-2 for the
XX PT treatment of hard tumors.
XX
XX PS Example 11; Page; 61pp; English.
XX
XX CC The invention relates to novel HER-2 (herstatin-2) antagonist
XX CC particularly a polypeptide that binds to the extracellular domain (ECD)
XX CC of HER-2 at a site that is different from the binding site of humanised
XX CC antibody, Herceptin, at an affinity of at least 108. The present
XX CC invention is based upon the initial discovery of an alternative HER-2
XX CC mRNA transcript with 274 bp insert of intron 8. The translation product
XX CC of the alternative transcript is a truncated HER-2 protein designated
XX CC p68HER-2 which lacks the transmembrane and intracellular domains of
XX CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX CC nucleic acids encoding these are useful to treat, diagnose and identify
XX CC solid tumours. The present sequence is p68HER-2 ECDIIIA variant encoded
XX CC by polymorphic form of human HER-2 intron 8. Note: The present sequence
XX CC is not shown in the specification but is derived from HER-2 intron 8
XX CC encoded ECDIIIA sequence given in figure 8 (AAE09184)
XX
XX SQ Sequence 79 AA;
XX
XX Query Match 96.2%; Score 402; DB 4; Length 79;
XX Best Local Similarity 96.2%; Pred. No. 1.2e-36;
XX Matches 76; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 GTHSLPPRPAAPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVSPVSV 60
XX Db 1 GTHSLPPRPAAPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVSPVSV 60
XX
XX QY 61 GRGPDPAHVAVNLSRYEG 79
XX Db 61 GRGPDPAHVAVNLSRYEG 79
XX
XX RESULT 13
XX AAE09189
XX ID AAE09189 standard; peptide; 79 AA.
XX
XX AC AAE09189;
XX
XX DT 15-NOV-2001 (first entry)
XX
XX DE Human p68HER-2 ECDIIIA variant 4 encoded by HER-2 intron 8.

```


XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
 XX ECDIIIA; variant.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 16
 FT /note= "p68HER-2 ECDIIIA (AAE09184) Leu substituted with
 FT Gln"
 XX
 FN WO2000161356-A1.
 XX
 PD 23-AUG-2001.
 XX
 PF 16-FEB-2001; 2001WO-US005327.
 XX
 PR 15-FEB-2000; 2000US-00506079.
 XX
 PA (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 PI Clinton G, Henner WD, Evans A;
 XX
 DR WPI; 2001-529934/58.
 DR N-PSDB; AAD15858.
 XX
 PT New polypeptide, which binds to the extracellular domain of HER-2 for the
 PT treatment of hard tumors.
 XX
 PS Example 11; Page; 61pp; English.
 XX
 CC The invention relates to novel HER-2 (herstatin-2) antagonist
 CC particularly a polypeptide that binds to the extracellular domain (ECD)
 CC of HER-2 at a site that is different from the binding site of humanised
 CC antibody, Herceptin, at an affinity of at least 10⁸. The present
 CC invention is based upon the initial discovery of an alternative HER-2
 CC mRNA transcript with 274 bp insert of intron 8. The translation product
 CC of the alternative transcript is a truncated HER-2 protein designated
 CC p68HER-2 which lacks the transmembrane and intracellular domains of
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
 CC nucleic acids encoding these are useful to treat, diagnose and identify
 CC solid tumours. The present sequence is p68HER-2 ECDIIIA variant encoded
 CC by polymorphic form of human HER-2 intron 8. Note: The present sequence
 CC is not shown in the specification but is derived from HER-2 intron 8
 CC encoded ECDIIIA sequence given in figure 8 (AAE09184)
 XX
 SQ Sequence 79 AA;

Query Match 95.7%; Score 400; DB 4; Length 79;
 Best Local Similarity 96.2%; Pred. No. 2.1e-36;
 Matches 76; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GTHTLLPRPAAPVPLRQPGPAHPVLFLRPSWDLVSAFYSLPLAPLSPTSPISPSV 60
 |||||
 Db 1 GTHTLLPRPAAPVPLRQPGPAHPVLFLRPSWDLVSAFYSLPLAPLSPTSPISPSV 60
 |||||

Qy 61 GRGPDPAHVAVNLSRYEG 79
 |||||
 Db 61 GRGPDPAHVAVNLSRYEG 79
 |||||

RESULT 14
 AAE09191
 ID AAE09191 standard; peptide; 79 AA.
 XX
 AC AAE09191;
 XX
 DT 15-NOV-2001 (first entry)
 XX
 DE Human p68HER-2 ECDIIIA variant 6 encoded by HER-2 intron 8.

XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
 XX ECDIIIA; variant.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 21
 FT /note= "p68HER-2 ECDIIIA (AAE09184) Gly substituted with
 FT Asp, Ala or Val"
 XX
 FN WO2000161356-A1.
 XX
 PD 23-AUG-2001.
 XX
 PF 16-FEB-2001; 2001WO-US005327.
 XX
 PR 16-FEB-2000; 2000US-00506079.
 XX
 PA (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 PI Clinton G, Henner WD, Evans A;
 XX
 DR WPI; 2001-529934/58.
 DR N-PSDB; AAD15860.
 XX
 PT New polypeptide, which binds to the extracellular domain of HER-2 for the
 PT treatment of hard tumors.
 XX
 PS Example 11; Page; 61pp; English.
 XX
 CC The invention relates to novel HER-2 (herstatin-2) antagonist
 CC particularly a polypeptide that binds to the extracellular domain (ECD)
 CC of HER-2 at a site that is different from the binding site of humanised
 CC antibody, Herceptin, at an affinity of at least 10⁸. The present
 CC invention is based upon the initial discovery of an alternative HER-2
 CC mRNA transcript with 274 bp insert of intron 8. The translation product
 CC of the alternative transcript is a truncated HER-2 protein designated
 CC p68HER-2 which lacks the transmembrane and intracellular domains of
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
 CC nucleic acids encoding these are useful to treat, diagnose and identify
 CC solid tumours. The present sequence is p68HER-2 ECDIIIA variant encoded
 CC by polymorphic form of human HER-2 intron 8. Note: The present sequence
 CC is not shown in the specification but is derived from HER-2 intron 8
 CC encoded ECDIIIA sequence given in figure 8 (AAE09184)
 XX
 SQ Sequence 79 AA;

Query Match 95.5%; Score 399; DB 4; Length 79;
 Best Local Similarity 96.2%; Pred. No. 2.7e-36;
 Matches 76; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GTHTLLPRPAAPVPLRQPGPAHPVLFLRPSWDLVSAFYSLPLAPLSPTSPISPSV 60
 |||||
 Db 1 GTHTLLPRPAAPVPLRQPGPAHPVLFLRPSWDLVSAFYSLPLAPLSPTSPISPSV 60
 |||||

Qy 61 GRGPDPAHVAVNLSRYEG 79
 |||||
 Db 61 GRGPDPAHVAVNLSRYEG 79
 |||||

RESULT 15
 AAE09187
 ID AAE09187 standard; peptide; 79 AA.
 XX
 AC AAE09187;
 XX
 DT 15-NOV-2001 (first entry)
 XX
 DE Human p68HER-2 ECDIIIA variant 2 encoded by HER-2 intron 8.

```

XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 5
FT /note= "p68HER-2 ECDIIIA (AAE09184) Leu substituted with
FT Pro"
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX N-PSDB; AAD15866.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX Example 11; Page; 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 108. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is p68HER-2 ECDIIIA variant encoded
XX by polymorphic form of human HER-2 intron 8. Note: The present sequence
XX is not shown in the specification but is derived from HER-2 intron 8
XX encoded ECDIIIA sequence given in figure 8 (AAE09184)
XX
XX SQ Sequence 79 AA;

Query Match 95.5%; Score 399; DB 4; Length 79;
Best Local Similarity 96.2%; Pred. No. 2.7e-36;
Matches 76; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTHSLPRAAIVPVLPMQPGPAHPVLSPFLRPSWDLVSAFYSLPLAPLSPTSPVSPVSV 60
Db 1 GTHSPFPRAAIVPVLPMQPGPAHPVLSPFLRPSWDLVSAFYSLPLAPLSPTSPVSPVSV 60

QY 61 GRGPDPAHVAVNLRYEG 79
Db 61 GRGPDPAHVAVDLSRYEG 79

RESULT 16
AAE09197
ID AAE09197 standard; peptide; 79 AA.
XX
XX AAE09197;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 ECDIIIA variant 13 encoded by HER-2 intron 8.

```

```

XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 17
FT /note= "p68HER-2 ECDIIIA (AAE09184) Arg substituted with
FT Cys"
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX N-PSDB; AAD15866.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX Example 12; Page; 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 108. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is p68HER-2 ECDIIIA variant encoded
XX by polymorphic form of human HER-2 intron 8. Note: The present sequence
XX is not shown in the specification but is derived from HER-2 intron 8
XX encoded ECDIIIA sequence given in figure 8 (AAE09184)
XX
XX SQ Sequence 79 AA;

Query Match 95.2%; Score 398; DB 4; Length 79;
Best Local Similarity 96.2%; Pred. No. 3.4e-36;
Matches 76; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTHSLPRAAIVPVLPMQPGPAHPVLSPFLRPSWDLVSAFYSLPLAPLSPTSPVSPVSV 60
Db 1 GTHSLPRAAIVPVLPMQPGPAHPVLSPFLRPSWDLVSAFYSLPLAPLSPTSPVSPVSV 60

QY 61 GRGPDPAHVAVNLRYEG 79
Db 61 GRGPDPAHVAVDLSRYEG 79

RESULT 17
AAE09198
ID AAE09198 standard; peptide; 79 AA.
XX
XX AAE09198;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 ECDIIIA variant 14 encoded by HER-2 intron 8.

```

```
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 31
FT /note= "p68HER-2 ECDIIIA (AAE09184) Arg substituted with
FT Ile"
XX
PN WO200161356-A1.
XX
PD 23-AUG-2001.
XX
PF 16-FEB-2001; 2001WO-US005327.
XX
PR 16-FEB-2000; 2000US-00506079.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Clinton G, Henner WD, Evans A;
XX
WPI; 2001-529934/58.
DR N-PSDB; AAD15867.
XX
PT New polypeptide, which binds to the extracellular domain of HER-2 for the
PT treatment of hard tumors.
XX
PS Example 12; Page; 61pp; English.
XX
CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 108. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is p68HER-2 ECDIIIA variant encoded
CC by polymorphic form of human HER-2 intron 8. Note: The present sequence
CC is not shown in the specification but is derived from HER-2 intron 8
CC encoded ECDIIIA sequence given in figure 8 (AAE09184)
XX
SQ Sequence 79 AA;
Query Match 95.2%; Score 398; DB 4; Length 79;
Best Local Similarity 96.2%; Pred. No. 3.4e-36;
Matches 76; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSFAFYSPLAPLSPTSPVPSV 60
DB 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSFAFYSPLAPLSPTSPVPSV 60
QY 61 GRGPDPAHVAVNLSRYEG 79
DB 61 GRGPDPAHVAVNLSRYEG 79
RESULT 18
AAE09193
ID AAE09193 standard; peptide; 79 AA.
XX
AC AAE09193;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 ECDIIIA variant 8 encoded by HER-2 intron 8.
```

```
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 54
FT /note= "p68HER-2 ECDIIIA (AAE09184) Pro substituted with
FT Arg"
XX
PN WO200161356-A1.
XX
PD 23-AUG-2001.
XX
PF 16-FEB-2001; 2001WO-US005327.
XX
PR 16-FEB-2000; 2000US-00506079.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Clinton G, Henner WD, Evans A;
XX
WPI; 2001-529934/58.
DR N-PSDB; AAD15862.
XX
PT New polypeptide, which binds to the extracellular domain of HER-2 for the
PT treatment of hard tumors.
XX
PS Example 11; Page; 61pp; English.
XX
CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 108. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is p68HER-2 ECDIIIA variant encoded
CC by polymorphic form of human HER-2 intron 8. Note: The present sequence
CC is not shown in the specification but is derived from HER-2 intron 8
CC encoded ECDIIIA sequence given in figure 8 (AAE09184)
XX
SQ Sequence 79 AA;
Query Match 95.0%; Score 397; DB 4; Length 79;
Best Local Similarity 96.2%; Pred. No. 4.4e-36;
Matches 76; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSFAFYSPLAPLSPTSPVPSV 60
DB 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSFAFYSPLAPLSPTSPVPSV 60
QY 61 GRGPDPAHVAVNLSRYEG 79
DB 61 GRGPDPAHVAVNLSRYEG 79
RESULT 19
AAE09194
ID AAE09194 standard; peptide; 79 AA.
XX
AC AAE09194;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 ECDIIIA variant 9 encoded by HER-2 intron 8.
```


CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIIA variant sequence. Note: The present sequence is not
CC shown in the specification but is derived from p68HER-2 generic sequence
CC (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX SQ Sequence 419 AA;

Query Match 86.8%; Score 363; DB 4; Length 419;
Best Local Similarity 88.6%; Pred. No. 1.7e-31;
Matches 70; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSMDLVSAFYSLPLAPLPTSPVPSVSV 60
Db 341 GXHSLPRPAAVPLRMQPGPAHPVLSFLRPSMDLVSAFYSLPLAPLPTSPVPSVSV 400

QY 61 GRGPDPAHVAVNLSRYEG 79
Db 401 GRGXDPDAHVAVNLSRYEG 419

RESULT 21
AAE09212
ID AAE09212 standard; protein; 419 AA.
XX AC AAE09212;
DT 15-NOV-2001 (first entry)
XX Human p68HER-2 generic protein variant 10.
DE HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX Homo sapiens.
FH Key Location/Qualifiers
FT Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124 /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125 /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= ECDIIIA variant
FT /note= "Extracellular domain IIIA variant"
FT Misc-difference 342 /label= Unknown
FT Misc-difference 345 /label= Unknown
FT Misc-difference 346 /label= Unknown
FT Misc-difference 356 /label= Unknown
FT Misc-difference 358 /label= Unknown
FT Misc-difference 361 /label= Unknown
FT Misc-difference 376 /label= Unknown
FT Misc-difference 394 /label= Unknown
FT Misc-difference 404 /label= Unknown
FT Misc-difference 413 /label= Unknown
FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
substituted with Asn"
XX WO200161356-A1.
XX 23-AUG-2001.
PD

XX 16-FEB-2001; 2001WO-US005327.
XX 16-FEB-2000; 2000US-00506079.
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX Clinton G, Henner WD, Evans A;
PI WPI; 2001-529934/58.
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX Example 11; Page; 61pp; English.
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 10⁻⁸. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence. Note: The present sequence is not
XX shown in the specification but is derived from p68HER-2 generic sequence
XX (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX SQ Sequence 419 AA;

Query Match 85.6%; Score 358; DB 4; Length 419;
Best Local Similarity 87.3%; Pred. No. 5.9e-31;
Matches 69; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSMDLVSAFYSLPLAPLPTSPVPSVSV 60
Db 341 GXHSLPRPAAVPLRMQPGPAHPVLSFLRPSMDLVSAFYSLPLAPLPTSPVPSVSV 400

QY 61 GRGPDPAHVAVNLSRYEG 79
Db 401 GRGXDPDAHVAVNLSRYEG 419

RESULT 22
AAE09205
ID AAE09205 standard; protein; 419 AA.
XX AC AAE09205;
DT 15-NOV-2001 (first entry)
XX Human p68HER-2 generic protein variant 3.
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
XX ECDIIIA; variant.
XX Homo sapiens.
FH Key Location/Qualifiers
FT Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124 /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125 /note= "Represented as Agn in the parent sequence shown
FT in the specification"

```

Domain      341.. 419
FT FT /label= ECDIIIA variant
FT FT /note= "Extracellular domain IIIA variant"
FT FT Misc-difference 342 /label= Unknown
FT FT Misc-difference 345 /label= Unknown
FT FT Misc-difference 346 /label= Unknown
FT FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT FT substituted with Leu"
FT FT Misc-difference 356 /label= Unknown
FT FT Misc-difference 358 /label= Unknown
FT FT Misc-difference 361 /label= Unknown
FT FT Misc-difference 376 /label= Unknown
FT FT Misc-difference 394 /label= Unknown
FT FT Misc-difference 404 /label= Unknown
FT FT Misc-difference 413 /label= Unknown
FT FT /label= Unknown
XX WO200161356-A1.
PN
XX
XX PD 23-AUG-2001.
XX
XX PF 16-FEB-2001; 2001WO-US005327.
XX
XX PR 16-FEB-2000; 2000US-00506079.
XX
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX PI Clinton G, Henner WD, Evans A;
XX
XX DR WPI; 2001-529934/58.
XX
XX PT New polypeptide, which binds to the extracellular domain of HER-2 for the
XX PT treatment of hard tumors.
XX PS Example 11; Page; 6pp; English.
XX
XX CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10-8. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIIA variant sequence. Note: The present sequence is not
CC shown in the specification but is derived from p68HER-2 generic sequence
CC (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
XX SQ Sequence 419 AA;
Query Match 85.2%; Score 356; DB 4; Length 419;
Best Local Similarity 87.3%; Pred.No.9.8e-31;
Matches 69; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 GTHTSLPRPAAPVPLRMQPGRPHPVLSFLRPSMDLVSAFYSFLAPLSPTSVPISFSV 60
DB 341 GXHXLPRPAAPVPPEXQXPAPHVLSFLRPSWDVXSFAFSLAPLDPTSVKISEVS 400
QY 61 GRGDEPDPAHVAVNLSRYEG 79

```

CC the ECD of HER-2 that is different from the site of binding for Herceptin
 CC (RTM) (a marketed humanized monoclonal antibody that is used for the
 CC treatment of cancer and binds to the ECD of HER-2). The methods,
 CC compositions, polypeptides and antibodies are used to treat solid tumours
 CC such as breast cancer, small cell lung carcinoma, ovarian cancer and/or
 CC colon cancer, especially where over-expression of HER-2 is indicated

XX Sequence 79 AA;

Query Match 84.9%; Score 355; DB 3; Length 79;
 Best Local Similarity 87.3%; Pred. No. 1.8e-31;
 Matches 69; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GTHSLPRPAAPVPLRMQPGPAHPVLSFLRPSWDILVSFAFSLPLAPLSPTSPVSPSV 60
 Db 1 GXHSXPRPAAPVPLRMQPGPAHPVLSFLRPSWDILVSFAFSLPLAPLSPTSPVSPSV 60

Qy 61 GRGPDPAHVAVNLRYEG 79
 Db 61 GRGPDPAHVAVNLRYEG 79

RESULT 24
 AAE09180
 ID AAE09180 standard; peptide; 79 AA.

AC AAE09180;

XX 15-NOV-2001 (first entry)

DE Human p68HER-2 ECDIIa generic sequence #1.

XX HER-2; herstatin; antagonist; extracellular domain; BCD; Herceptin;
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
 KW ECDIIa.

XX Homo sapiens.

Key Location/Qualifiers
 FT Misc-difference 2 /label= Unknown
 FT /note= "Encoded by WCC"
 FT Misc-difference 5 /label= Unknown
 FT /note= "Encoded by CYG"
 FT Misc-difference 6 /label= Unknown
 FT /note= "Encoded by CYC"
 FT Misc-difference 16 /label= Unknown
 FT /note= "Encoded by CWG"
 FT Misc-difference 18 /label= Unknown
 FT /note= "Encoded by ATG"
 FT Misc-difference 21 /label= Unknown
 FT /note= "Encoded by GNC"
 FT Misc-difference 36 /label= Unknown
 FT /note= "Encoded by MTA"
 FT Misc-difference 34 /label= Unknown
 FT /note= "Encoded by CST"
 FT Misc-difference 64 /label= Unknown
 FT /note= "Encoded by CYG"
 FT Misc-difference 73 /label= Unknown
 FT /note= "Encoded by SAC"

XX WO200161356-A1.

XX 23-AUG-2001.

XX 16-FEB-2001; 2001WO-US005327.
 XX 16-FEB-2000; 2000US-00506079.
 XX (UYOR-) UNIV OREGON HEALTH SCI.

PI Clinton G, Henner WD, Evans A;

DR WPI; 2001-529934/58.
 DR N-PSDB; AAD15844.

XX New polypeptide, which binds to the extracellular domain of HER-2 for the
 PT treatment of hard tumors.

PS Claim 1; Page 52-53; 61pp; English.

XX The invention relates to novel HER-2 (herstatin-2) antagonist
 CC particularly a polypeptide that binds to the extracellular domain (ECD)
 CC of HER-2 at a site that is different from the binding site of humanised
 CC antibody, Herceptin, at an affinity of at least 10⁸. The present
 CC invention is based upon the initial discovery of an alternative HER-2
 CC mRNA transcript with 274 bp insert of intron 8. The translation product
 CC of the alternative transcript is a truncated HER-2 protein designated
 CC p68HER-2 which lacks the transmembrane and intracellular domains of
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIa.
 CC The ECDIIa-containing polypeptides bind tightly to, and thus antagonise
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
 CC nucleic acids encoding these are useful to treat, diagnose and identify
 CC solid tumours. The present sequence is human p68HER-2 ECDIIa peptide
 CC generic sequence

XX Sequence 79 AA;

Query Match 84.9%; Score 355; DB 4; Length 79;
 Best Local Similarity 87.3%; Pred. No. 1.8e-31;
 Matches 69; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GTHSLPRPAAPVPLRMQPGPAHPVLSFLRPSWDILVSFAFSLPLAPLSPTSPVSPSV 60
 Db 1 GXHSXPRPAAPVPLRMQPGPAHPVLSFLRPSWDILVSFAFSLPLAPLSPTSPVSPSV 60

Qy 61 GRGPDPAHVAVNLRYEG 79
 Db 61 GRGPDPAHVAVNLRYEG 79

RESULT 25
 AAE20347
 ID AAE20347 standard; protein; 79 AA.

XX AAE20347;

XX 18-JUN-2002 (first entry)

XX Human HER2 intron 8 polymorphic sequence encoded protein.

XX Human; tumour; endothelial growth factor receptor; EGFR; cytostatic;
 KW herstatin; HER-2 receptor tyrosine kinase; squamous cell carcinoma; lungs;
 KW colon; glial cell tumour; cell growth; polymorphism.

XX Homo sapiens.

Key Location/Qualifiers
 FT Misc-difference 2 /label= Thr, Ser
 FT /note= "Encoded by WCC"
 FT Misc-difference 5 /label= Leu, Pro
 FT /note= "Encoded by CYG"
 FT Misc-difference 6 /label= Pro, Leu
 FT /note= "Encoded by CYC"

FT Misc-difference 16 /label= Leu, Gln
 FT /note= "Encoded by CWG"
 FT Misc-difference 18 /label= Met, Leu
 FT /note= "Encoded by ATR"
 FT Misc-difference 21 /label= Gly, Asp, Ala, Val
 FT /note= "Encoded by GNC"
 FT Misc-difference 36 /label= Leu, Ile
 FT /note= "Encoded by MTA"
 FT Misc-difference 54 /label= Pro, Arg
 FT /note= "Encoded by CST"
 FT Misc-difference 64 /label= Pro, Leu
 FT /note= "Encoded by CYG"
 FT Misc-difference 73 /label= Asp, Asn
 FT /note= "Encoded by SAC"
 FT
 FT WO200214470-A2.
 FT
 FT 21-FEB-2002.
 FT
 FT 14-AUG-2001; 2001WO-US025502.
 FT
 FT 14-AUG-2000; 2000US-00638834.
 FT
 FT (UYOR-) UNIV OREGON HEALTH SCI.
 FT
 FT Clinton GW;
 FT
 FT WPI; 2002-269185/31.
 FT N-ESDB; AAD32538.
 FT
 FT Treating solid tumor characterized by expression of endothelial growth
 FT factor receptor, involves administering recombinant herstatin that binds
 FT to extracellular domain of the endothelial growth factor receptor.
 FT
 FT Claim 1; Page 77; 82pp; English.
 FT
 FT The present invention relates to a method for treating a solid tumor
 FT characterised by endothelial growth factor receptor (EGFR) expression.
 FT The method involves administering an agent that binds to an extracellular
 FT domain (ECD) of EGFR. The invention also relates to a naturally occurring
 FT inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-
 FT expression of herstatin with p18HER2 causes a striking reduction in cell
 FT growth that corresponds with suppression of p185 autophosphorylation. The
 FT method or a pharmaceutical composition is useful for treating a solid
 FT tumour (selected from squamous cell carcinoma, lung carcinoma, colon
 FT carcinoma and glial cell tumour) characterised by EGFR expression. The
 FT present sequence is a protein encoded by human HER2 intron 8 polymorphic
 FT sequence
 FT
 FT Sequence 79 AA;
 FT
 FT Query Match 84.9%; Score 355; DB 5; Length 79;
 FT Best Local Similarity 87.3%; Pred. No. 1.8e-31;
 FT Matches 69; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 FT
 FT 1 GTHSILPEPAVYPLRMQGPAPHVLSFLRPSWDLVSFAFYSLEPLAPLSPTSVISPSV 60
 FT
 FT 1 GXHXXPRPAVAVPVRXQRPAPHVLSFLRPSWDLVSFAFYSLEPLAPLSPTSVISPSV 60
 FT
 FT 61 GRGPDPAHVAVNLSRYEG 79
 FT
 FT 61 GRGXDPDAHVAVXLSRYEG 79
 FT
 FT RESULT 26
 FT AAE09207

ID XX AAE09207 standard; protein; 419 AA.
 XX AAE09207;
 XX
 DT 15-NOV-2001 (first entry)
 XX
 DE Human p68HER-2 generic protein variant 5.
 XX
 KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
 KW ECDIIIA; variant.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..340
 FT /note= "Identical to N-terminal region of p185HER-2"
 FT Misc-difference 124 /note= "Represented as Agn in the parent sequence shown
 FT in the specification"
 FT Misc-difference 125 /note= "Represented as Agn in the parent sequence shown
 FT in the specification"
 FT Domain 341..419
 FT /label= ECDIIIA variant
 FT /note= "Extracellular domain IIIa variant"
 FT Misc-difference 342 /label= Unknown
 FT Misc-difference 345 /label= Unknown
 FT Misc-difference 346 /label= Unknown
 FT Misc-difference 356 /label= Unknown
 FT Misc-difference 358 /label= Unknown
 FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
 FT substituted with Leu"
 FT Misc-difference 361 /label= Unknown
 FT Misc-difference 376 /label= Unknown
 FT Misc-difference 394 /label= Unknown
 FT Misc-difference 404 /label= Unknown
 FT Misc-difference 413 /label= Unknown
 FT
 XX WO200161356-A1.
 XX
 XX 23-AUG-2001.
 XX
 XX 16-FEB-2001; 2001WO-US005327.
 XX
 XX 16-FEB-2000; 2000US-00506079.
 XX
 XX (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 XX Clinton G, Henner WD, Evans A;
 XX WPI; 2001-529934/58.
 XX
 XX New polypeptide, which binds to the extracellular domain of HER-2 for the
 XX treatment of hard tumors.
 XX
 XX Example 11; Page; 61pp; English.
 XX
 XX The invention relates to novel HER-2 (herstatin-2) antagonist
 XX particularly a polypeptide that binds to the extracellular domain (ECD)
 XX of HER-2 at a site that is different from the binding site of humanised
 XX antibody, Herceptin, at an affinity of at least 10⁶. The present
 XX invention is based upon the initial discovery of an alternative HER-2
 XX mRNA transcript with 274 bp insert of intron 8. The translation product

CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIIA variant sequence. Note: The present sequence is not
CC shown in the specification but is derived from p68HER-2 generic sequence
CC (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
SQ Sequence 419 AA;

Query Match 84.7%; Score 354; DB 4; Length 419;
Best Local Similarity 86.1%; Pred. No. 1.6e-30;
Matches 68; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
QY 1 GTHSLLPRAVPVPLRMQPGAPVLSFLRPSMDLVSFAFVSLPLAPLSPTSPVSPVSV 60
DB 341 GXHSXXPRPAVFPVXLQXPAPHPVLSFLRPSMDLVSFAFVSLPLAPLSPTSPVSV 400
QY 61 GRGPDPAHVAVNLSRYEG 79
DB 401 GRGPDPAHVAVNLSRYEG 419

RESULT 27
AAE09209
ID AAE09209 standard; protein; 419 AA.
XX AC AAE09209;
XX 15-NOV-2001 (first entry)
DE Human p68HER-2 generic protein variant 7.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
OS Homo sapiens.

XX Key Location/Qualifiers
FT Region 1..340
FT Misc-difference 124 /note= "Identical to N-terminal region of p185HER-2"
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125 /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= ECDIIIA variant
FT /note= "Extracellular domain IIIA variant"
FT Misc-difference 342 /label= Unknown
FT Misc-difference 345 /label= Unknown
FT Misc-difference 346 /label= Unknown
FT Misc-difference 356 /label= Unknown
FT Misc-difference 358 /label= Unknown
FT Misc-difference 361 /label= Unknown
FT Misc-difference 376 /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Ile"
FT Misc-difference 394 /label= Unknown
FT Misc-difference 404 /label= Unknown
FT /label= Unknown

FT Misc-difference 413 /label= Unknown
FT WO200161356-A1.
XX 23-AUG-2001.
PD
XX 16-FEB-2001; 2001WO-US005327.
XX 16-FEB-2000; 2000US-00506079.
PR
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
PI WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX Example 11; Page; 61pp; English.

XX The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10⁻⁸. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIIA variant sequence. Note: The present sequence is not
CC shown in the specification but is derived from p68HER-2 generic sequence
CC (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
SQ Sequence 419 AA;

Query Match 84.7%; Score 354; DB 4; Length 419;
Best Local Similarity 86.1%; Pred. No. 1.6e-30;
Matches 68; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
QY 1 GTHSLLPRAVPVPLRMQPGAPVLSFLRPSMDLVSFAFVSLPLAPLSPTSPVSPVSV 60
DB 341 GXHSXXPRPAVFPVXLQXPAPHPVLSFLRPSMDLVSFAFVSLPLAPLSPTSPVSV 400
QY 61 GRGPDPAHVAVNLSRYEG 79
DB 401 GRGPDPAHVAVNLSRYEG 419

RESULT 28
AAE09203
ID AAE09203 standard; protein; 419 AA.
XX AC AAE09203;
XX 15-NOV-2001 (first entry)
XX Human p68HER-2 generic protein variant 1.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
XX Homo sapiens.
XX Key Location/Qualifiers
FT Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"

```

FT Misc-difference 124 /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125 /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= "ECDIIIIa variant"
FT /note= "Extracellular domain IIIa variant"
FT Misc-difference 342 /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Ser"
FT Misc-difference 345 /label= Unknown
FT Misc-difference 346 /label= Unknown
FT Misc-difference 356 /label= Unknown
FT Misc-difference 358 /label= Unknown
FT Misc-difference 361 /label= Unknown
FT Misc-difference 376 /label= Unknown
FT Misc-difference 394 /label= Unknown
FT Misc-difference 404 /label= Unknown
FT Misc-difference 413 /label= Unknown
FT
FT
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX {UYOR-} UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX Example 11; Page; 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 108. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIIa.
XX The ECDIIIIa-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIIa variant sequence. Note: The present sequence is not
XX shown in the specification but is derived from p68HER-2 generic sequence
XX (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
XX Sequence 419 AA;
XX
XX Query Match 84.2%; Score 352; DB 4; Length 419;
XX Best Local Similarity 86.1%; Pred. No. 2.7e-30;
XX Matches 68; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

```

```

Qy 1 GTHSLPRPAAPVPLRMQGPAPVLSFLPSKWDLVSAFYSLPLAPLSPTSVSPVSV 60
Db 341 GSHSXXPRPAAPVVPXEXQXPAPVLSFLPSKWDLVSAFYSLPLAPLSPTSVSPVSV 400
Qy 61 GRGPDPPDAHVAVNLSRYEG 79
Db 401 GRGXDPDAHVAVNLSRYEG 419

RESULT 29
AAE09210
ID AAE09210 standard; protein; 419 AA.
AC AAE09210;
XX
XX 15-NOV-2001 (first entry)
DT Human p68HER-2 generic protein variant 8.
DE
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
XX ECDIIIIa; variant.
XX
XX Homo sapiens.
XX
XX OS
XX
XX Key Location/Qualifiers
XX Region 1..340
XX /note= "Identical to N-terminal region of p185HER-2"
XX Misc-difference 124
XX /note= "Represented as Agn in the parent sequence shown
XX in the specification"
XX Misc-difference 125
XX /note= "Represented as Agn in the parent sequence shown
XX in the specification"
XX Domain 341..419
XX /label= "ECDIIIIa variant"
XX /note= "Extracellular domain IIIa variant"
XX Misc-difference 342
XX /label= Unknown
XX Misc-difference 345
XX /label= Unknown
XX Misc-difference 346
XX /label= Unknown
XX Misc-difference 356
XX /label= Unknown
XX Misc-difference 358
XX /label= Unknown
XX Misc-difference 361
XX /label= Unknown
XX Misc-difference 376
XX /label= Unknown
XX Misc-difference 394
XX /note= "p68HER-2 generic sequence (AAE09181) Xaa
XX substituted with Arg"
XX Misc-difference 404
XX /label= Unknown
XX Misc-difference 413
XX /label= Unknown
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX {UYOR-} UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX

```

PT New polypeptide, which binds to the extracellular domain of HER-2 for the
PT treatment of hard tumors.
XX
PS Example 11; Page; 61pp; English.
XX
CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10⁻⁸. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC cDNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIIA variant sequence. Note: The present sequence is not
CC shown in the specification but is derived from p68HER-2 generic sequence
CC (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
SQ Sequence 419 AA;

Query Match 84.0%; Score 351; DB 4; Length 419;
Best Local Similarity 86.1%; Pred. No. 3.5e-30;
Matches 68; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 GTHSLIPRAAVVPLRMQPGPAHFVLSFLRPSWDLVSFAFYSPLAPLPTSVIPSPVSV 60
DB 341 GXHSXXPRPAAVVPEXQXPFAHFVLSFLRPSWDXVSFAFYSPLAPLPTSVRISPSVSV 400

QY 61 GRGPPDDAHVAVNLSRYEG 79
DB 401 GRGXDPDAHVAVXLSRYEG 419

RESULT 30
AAE09208
ID AA209208 standard; protein; 419 AA.
AC AA209208;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 generic protein variant 6.
XX
KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumor; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= ECDIIIA variant
FT /note= "Extracellular domain IIIa variant"
FT Misc-difference 342
FT /label= Unknown
FT Misc-difference 345
FT /label= Unknown
FT Misc-difference 346
FT /label= Unknown
FT Misc-difference 356
FT /label= Unknown

FT Misc-difference 358
FT /label= Unknown
FT Misc-difference 361
FT /label= Asp, Ala, Val
FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown
XX
PW WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX Example 11; Page; 61pp; English.
XX
CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10⁻⁸. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIIA variant sequence. Note: The present sequence is not
CC shown in the specification but is derived from p68HER-2 generic sequence
CC (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
SQ Sequence 419 AA;

Query Match 84.0%; Score 351; DB 4; Length 419;
Best Local Similarity 86.1%; Pred. No. 3.5e-30;
Matches 68; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 GTHSLIPRAAVVPLRMQPGPAHFVLSFLRPSWDLVSFAFYSPLAPLPTSVIPSPVSV 60
DB 341 GXHSXXPRPAAVVPEXQXPFAHFVLSFLRPSWDXVSFAFYSPLAPLPTSVXISPSVSV 400

QY 61 GRGPPDDAHVAVNLSRYEG 79
DB 401 GRGXDPDAHVAVXLSRYEG 419

Search completed: July 4, 2004, 04:18:46
Job time : 20.3394 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2004, 04:18:59 ; Search time 5.23494 Seconds
(without alignments)
779.083 Million cell updates/sec

Title: US-09-506-079H-11
Perfect score: 418
Sequence: 1 GTHSLPRPAVPLRMPQ.....VGRGPDPAHVNLRYEG 79

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 90 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/2/iaa/5A-COMB.pep:*

2: /cgn2_6/prodata/2/iaa/5B-COMB.pep:*

3: /cgn2_6/prodata/2/iaa/6A-COMB.pep:*

4: /cgn2_6/prodata/2/iaa/6B-COMB.pep:*

5: /cgn2_6/prodata/2/iaa/ECTUS-COMB.pep:*

6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	418	100.0	79	4	US-09-630-155-1
2	418	100.0	419	4	US-09-630-155-2
3	83.5	20.0	888	4	US-09-077-940A-4
4	81.5	19.5	1257	1	US-08-340-428B-49
5	81	19.4	158	4	US-09-252-991A-24956
6	79	18.9	195	4	US-09-252-991A-29314
7	77	18.4	122	4	US-09-462-606-65
8	77	18.4	546	4	US-09-252-991A-19122
9	74	17.7	495	4	US-09-252-991A-31949
10	73.5	17.6	887	4	US-09-077-940A-2
11	72	17.2	183	4	US-09-489-039A-12842
12	72	17.2	203	4	US-09-543-681A-8287
13	72	17.2	440	3	US-08-430-286A-9
14	71.5	17.1	604	2	US-08-468-578B-12
15	71.5	17.1	604	3	US-08-468-579B-12
16	71.5	17.1	604	3	US-08-468-577B-12
17	71.5	17.1	2441	1	US-08-194-468-2
18	71.5	17.1	2441	3	US-08-961-739-2
19	71.5	17.1	2441	4	US-09-514-247A-8
20	71.5	17.1	2441	4	US-09-686-316-2
21	71	17.0	803	4	US-09-252-991A-30479
22	70	16.7	2321	4	US-09-230-652-2
23	69.5	16.6	344	4	US-09-147-236-11
24	69.5	16.6	344	4	US-09-522-474-11
25	69.5	16.6	432	1	US-08-615-170-21
26	69.5	16.6	433	1	US-08-615-170-19
27	68.5	16.4	115	3	US-09-461-697-58

ALIGNMENTS

RESULT 1
US-09-630-155-1
; Sequence 1, Application US/09630155
; Patent No. 6414130
; GENERAL INFORMATION:

APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVIS WRIGHT TREMAINE LLP
STREET: 1501 Fourth Avenue, 2600 Century Square
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: PC compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Word

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/630,155
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Davison, Barry L.
REGISTRATION NUMBER: 47,309
REFERENCE/DOCKET NUMBER: 49321-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 628-7621
TELEFAX: 206 628-7699

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 79

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: HER-2 ECD antagonist
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-630-155-1

Query Match 100.0%; Score 418; DB 4; Length 79;
Best Local Similarity 100.0%; Pred. No. 8.7e-42;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSVPIPSVSV 60
DB 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSVPIPSVSV 60
QY 61 GRGPDPAHVAVNLSRYEG 79
DB 61 GRGPDPAHVAVNLSRYEG 79

RESULT 2
US-09-630-155-2
Sequence 2, Application US/09630155
Patent No. 641430

GENERAL INFORMATION:
APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVIS WRIGHT TREMAINE LLP
STREET: 1501 Fourth Avenue, 2600 Century Square
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: PC compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Word

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/630,155
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Davison, Barry L.
REGISTRATION NUMBER: 47,309
REFERENCE/DOCKET NUMBER: 49321-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 628-7621
TELEFAX: 206 628-7699
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: polypeptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-630-155-2

Query Match 100.0%; Score 418; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 7.1e-41;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSVPIPSVSV 60
DB 341 GTHSLPRPAAVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSVPIPSVSV 400
QY 61 GRGPDPAHVAVNLSRYEG 79
DB 401 GRGPDPAHVAVNLSRYEG 419

RESULT 3
US-09-077-940A-4
Sequence 4, Application US/09077940A
Patent No. 6576441

GENERAL INFORMATION:
APPLICANT: KIMURA, Toru et al.
TITLE OF INVENTION: NOVEL SEMAPHORIN Z AND GENE ENCODING THE SAME
FILE REFERENCE: 0020-4426P
CURRENT APPLICATION NUMBER: US/09/077,940A
CURRENT FILING DATE: 1998-06-05
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patent in version 3.1
SEQ ID NO 4
LENGTH: 888
TYPE: PRT
ORGANISM: Homo sapiens
US-09-077-940A-4

Query Match 20.0%; Score 83.5; DB 4; Length 888;
Best Local Similarity 35.4%; Pred. No. 0.22;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

QY 1 GTH----SLLPRPAAVPLRMQPGPAHPVLSFLRP-SWD---LVSAFYSLPLAPLSPT 51
DB 698 GPHDLDSGLLPTPEQTPLPQRLPTP-HPHPHALGPRANDHGHPLLPASASSSLLLLAPA 756
QY 52 SVPIPSVSVGRGPDPAHV 70
DB 757 RAPEQPPAPGE-PTFDGRL 774

RESULT 4
US-08-340-428B-49
Sequence 49, Application US/08340428B
Patent No. 5648465

GENERAL INFORMATION:
APPLICANT: MARGOLIS, Richard U.
APPELLANT: RAUCH, Uwe
APPELLANT: MARGOLIS, Renee K.
TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A
TITLE OF INVENTION: NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:

ADDRESS: Browdy and Neimark
STREET: 419 Seventh Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340.428B
FILING DATE: 14 No. 5648465ember 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/922.911
FILING DATE: 03 August 1992
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25.618
REFERENCE/DOCKET NUMBER: Margolis-1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 1257 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-340-428B-49

Query Match 19.5%; Score 81.5; DB 1; Length 1257;
Best Local Similarity 35.0%; Pred. No. 0.58; Indels 17; Gaps 5;
Matches 28; Conservative 10; Mismatches 25
Qy 4 SLLPRAAVPVPLRMQPG---PAHVLVSLFRL-----PSWDLVSFAFVSLPLAPLS--PT 51
Db 610 SSIPSEALSAVSLQASPGDGDFFIVAMLRAPKLLPHSTLVNVPSPILSPASPLPS 669
Qy 52 SVP-----ISPVSVGRGDPDP 67
Db 670 SVPERQAVRVSFG-AEDPE 688

RESULT 5
US-09-252-991A-24956
; Sequence 24956, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24956
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24956

Query Match 19.4%; Score 81; DB 4; Length 158;
Best Local Similarity 36.2%; Pred. No. 0.049; Indels 8; Gaps 3;
Matches 25; Conservative 8; Mismatches 28

Qy 6 LPRP-----AAYPVPLRMQPGPAHVLVSLFRLPSWDLVSFAFVSLPL-APLSPTSPVSPV 58
Db 87 LPSPPPPPPPFLSLPLFLPLFLPLFLPLFLPLFLPLFLPLFLPLFLPLFLPLFLPLPL 146
Qy 59 SV-GRGDP 66
Db 147 SLSSSSPSP 155

RESULT 6
US-09-252-991A-29314
; Sequence 29314, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29314
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29314

Query Match 18.9%; Score 79; DB 4; Length 195;
Best Local Similarity 37.9%; Pred. No. 0.11; Indels 10; Gaps 3;
Matches 25; Conservative 1; Mismatches 30
Qy 1 GTHSLPRAAVPVPLRMQPGPAHVLVSLFRLPSWDLVSFAFVSLPLAPLSPTSPVSPV 60
Db 19 GLRHTAPRAAVPDP---QPA-RHGPVVRP-----RAARQAPAGPALPRTTLTLPORH 68
Qy 61 GRGDP 66
Db 69 GRQDP 74

RESULT 7
US-09-462-606-65
; Sequence 65, Application US/09462606
; Patent No. 6432408
; GENERAL INFORMATION:
; APPLICANT: MENG, XIANG-JIN
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: A SWINE HEPATITIS E VIRUS AND USES THEREOF
; FILE REFERENCE: 20264267US1
; CURRENT APPLICATION NUMBER: US/09/462,606
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: US 60/053069
; PRIOR FILING DATE: 1997-07-18
; PRIOR APPLICATION NUMBER: PCT/US98/14665
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 65
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Hepatitis E virus
US-09-462-606-65

Query Match 18.4%; Score 77; DB 4; Length 122;
Best Local Similarity 36.1%; Pred. No. 0.1; Indels 6; Gaps 3;
Matches 22; Conservative 10; Mismatches 23

QY 1 GTHSLPRPAVPVPLRMOPGPAHPVLSFLRPSNDLVSAFYSLPLAPL---SPTSVPISP 57
DB 55 GVTGLILSPS--PSPIFIQPTPSLP-MSFHNFGLEFALDSRPAPLAPLGVTSAPLPP 111
QY 58 V 58
DB 112 V 112

RESULT 8
US-09-252-991A-19122
; Sequence 19122, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19122
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19122

Query Match 18.4%; Score 77; DB 4; Length 546;
Best Local Similarity 31.2%; Pred. No. 0.69;
Matches 24; Conservative 14; Mismatches 25; Indels 14; Gaps 4;

QY 3 HSLPRPAVPVPLRMOPGPAHPVLSFLRPSNDLVSAFYSLPLAPLSPVSVG 61
DB 42 HALVGRGQLPRLRPF--YPVAL-----QVGLLQLPARPVAPGAVQARILRPG 92
QY 62 R---GPDPAHVAVNL 74
DB 93 RGLLAGADHEAHLGVH 109

RESULT 9
US-09-252-991A-31949
; Sequence 31949, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31949
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31949

Query Match 17.7%; Score 74; DB 4; Length 495;
Best Local Similarity 39.1%; Pred. No. 1.4;
Matches 25; Conservative 9; Mismatches 24; Indels 24; Gaps 4;

QY 8 RPAVPVPLRMOPGPAHPVLSFL--RPSNDLVSAFYSLPLAPLSPVSVG 65
DB 284 RPAQPRE-GAGPGPARPVATGVARRAGDALA--SARPARI-PTTILRSAPDAAPGA 339

QY 66 PDAH 69
DB 340 PIPH 343

RESULT 10
US-09-077-940A-2
; Sequence 2, Application US/09077940A
; Patent No. 6576441
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru et al.
; TITLE OF INVENTION: NOVEL SEMAPHORIN 2 AND GENE ENCODING THE SAME
; FILE REFERENCE: 0020-4426P
; CURRENT APPLICATION NUMBER: US/09/077,940A
; CURRENT FILING DATE: 1998-06-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 887
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-077-940A-2

Query Match 17.6%; Score 73.5; DB 4; Length 887;
Best Local Similarity 36.0%; Pred. No. 3.2;
Matches 27; Conservative 6; Mismatches 31; Indels 11; Gaps 5;

QY 1 GTH---SLLPRPAVPVPLRMOPGPAHPVLSFLRSP-SWD----LVSAFYSLPLAPLSP 51
DB 699 GHDLDSGLLPTPEQTPLPKSLP-TTHPHAHGPRAWDHSALLSASASTSLLLAHT 757
QY 52 SVPIIS-PVSVGRGPD 65
DB 758 RAPEQPPVPTESGPE 772

RESULT 11
US-09-489-039A-12842
; Sequence 12842, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12842
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12842

Query Match 17.2%; Score 72; DB 4; Length 183;
Best Local Similarity 23.8%; Pred. No. 0.66;
Matches 30; Conservative 8; Mismatches 30; Indels 58; Gaps 5;

QY 4 SLLPRPAVPVPL-----RMQP-----GPAH-PVLSFLRPS 33
DB 23 SSAPRPQLSPQLPPVSSPRQLSLQRLRPASSRQPPWPPVPLPLSPGHPPLSPRRPL 82
QY 34 MDLVSAFYSLPLAPL-----SPTSVPISPVSVGRGPD 66
DB 83 WQRLPALL-LPYRFLMLLPFRPWQPLLPASAPSLSLPLPAPAAAPVRAVSAVPVPP 141
QY 67 DAHVAV 72
DB 142 VSHVTL 147

```
RESULT 12
US-09-543-681A-8287
; Sequence 8287, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRITON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 8287
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (14), (15), (16), (17), (18), (19), (20), (21), (22), (23), (24), (25), (26)
; LOCATION: (27), (28), (29), (30), (31), (32), (33), (34), (35), (36), (37), (38), (39)
; LOCATION: (40), (41), (42), (43), (44), (45), (46), (47), (48), (49), (50), (51), (52)
; LOCATION: (53), (54), (55), (56), (57), (58), (59), (60), (61), (62), (63), (72), (78)
; LOCATION: (79), (83), (90), (103)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-543-681A-8287

Query Match 17.2%; Score 72; DB 4; Length 203;
Best Local Similarity 35.8%; Pred. No. 0.76;
Matches 19; Conservative 4; Mismatches 26; Indels 4; Gaps 1;

QY 7 PREAAVPELRMQPGAPHVLSFLRPSWDLVSFAFSLPLAPLSPTSPVPSVS 59
DB 91 PEPPEPPPPPSXPPPPHPLPLSP-----LFFYIYLFHPIYILSPFIFSS 139

RESULT 13
US-08-430-286A-9
; Sequence 9, Application US/08430286A
; Patent No. 6225080
; GENERAL INFORMATION:
; APPLICANT: Uhl, George R.
; APPLICANT: Eppler, C. Mark
; APPLICANT: Wang, Jai-Bel
; TITLE OF INVENTION: Mu-Subtype Opioid Receptor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/430,286A
; FILING DATE: 28-APR-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Robinson, Joseph R.
; REGISTRATION NUMBER: 33,448
; REFERENCE/DOCKET NUMBER: 0646/LA843-US5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7700
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: OPB-R
US-08-430-286A-9

Query Match 17.2%; Score 72; DB 3; Length 440;
Best Local Similarity 45.2%; Pred. No. 2;
Matches 19; Conservative 5; Mismatches 10; Indels 8; Gaps 3;

QY 2 THSLLRPAAPVPELRMQPGAPHVLSFLRPSWDLVSFAFYSL 43
DB 32 TASPSPAPSWTPSP---RPGPAHP---FLOPPWAV--ALMSL 65

RESULT 14
US-08-468-576B-12
; Sequence 12, Application US/08468576B
; Patent No. 5955345
; GENERAL INFORMATION:
; APPLICANT: Rabin, Daniel
; TITLE OF INVENTION: PANCREATIC ISLET CELL ANTIGENS
; TITLE OF INVENTION: OBTAINED BY MOLECULAR CLONING
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sprung Kramer Schaefer & Briscoe
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-5144
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.5
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,576B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/239,276
; FILING DATE: 05-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/872,646
; FILING DATE: 08-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/715,181
; FILING DATE: 14-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/441,703
; FILING DATE: 04-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/312,543
; FILING DATE: 17-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: MDI 251.7-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 604 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```


Db 480 ILSSLN 485

RESULT 17

US-08-194-468-2

Sequence 2, Application US/08194468

Patent No. 5250336

GENERAL INFORMATION:

APPLICANT: Montminy, Marc R.

TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN

TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN

TITLE OF INVENTION: RESPONSIVE GENES

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Pretty, Schroeder, Bruggemann & Clark

STREET: 444 South Flower Street, Suite 2000

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/194,468

FILING DATE: 10-FEB-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Reiter, Stephen E.

REGISTRATION NUMBER: 31,192

REFERENCE/DOCKET NUMBER: P41 9672

TELEPHONE: (619) 546-4737

TELEFAX: (619) 546-9392

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2441 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-194-468-2

Query Match 17.1%; Score 71.5; DB 1; Length 2441;

Best Local Similarity 36.5%; Pred. No. 20;

Matches 23; Conservative 4; Mismatches 33; Indels 3; Gaps 2;

Qy 4 SLLPRPAVPVPLRMQGPAPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVPSVSGRG 63

Db 843 SOLPCPPVTQSLHPTPPASTAAGM--PSLQHPATPGMTPPQPAATQ-PSTPVSSGQT 899

Qy 64 PDP 66

Db 900 PTP 902

RESULT 18

US-08-961-739-2

Sequence 2, Application US/08961739A

Patent No. 6063583

GENERAL INFORMATION:

APPLICANT: Montminy, Marc R.

TITLE OF INVENTION: Methods for Treating Diabetes Mellitus

FILE REFERENCE: SALK1650-1

CURRENT APPLICATION NUMBER: US/08/961,739A

CURRENT FILING DATE: 1997-10-31

EARLIER APPLICATION NUMBER: US 194,468

EARLIER FILING DATE: 1994-02-10

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 2441

; TYPE: PRT

; ORGANISM: Mus

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: {1}...(2441)

; OTHER INFORMATION: Xaa = Any Amino Acid

US-08-961-739-2

Query Match 17.1%; Score 71.5; DB 3; Length 2441;

Best Local Similarity 36.5%; Pred. No. 20;

Matches 23; Conservative 4; Mismatches 33; Indels 3; Gaps 2;

Qy 4 SLLPRPAVPVPLRMQGPAPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVPSVSGRG 63

Db 843 SOLPCPPVTQSLHPTPPASTAAGM--PSLQHPATPGMTPPQPAATQ-PSTPVSSGQT 899

Qy 64 PDP 66

Db 900 PTP 902

RESULT 19

US-09-514-247A-8

Sequence 8, Application US/09514247A

Patent No. 6365361

GENERAL INFORMATION:

APPLICANT: TANABE SEIYAKU CO. LTD.

APPLICANT: TANIGUCHI, Tomoyasu

APPLICANT: MIZUKAMI, Junko

TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO P

FILE REFERENCE: TANIGUCHI-6

CURRENT APPLICATION NUMBER: US/09/514,247A

CURRENT FILING DATE: 2000-02-28

PRIOR APPLICATION NUMBER: PCT/JP98/03734

PRIOR FILING DATE: 1998-08-24

PRIOR APPLICATION NUMBER: JP231084/1997

PRIOR FILING DATE: 1997-08-27

NUMBER OF SEQ ID NOS: 10

SOFTWARE: PatentIn version 3.0

SEQ ID NO 8

LENGTH: 2441

TYPE: PRT

; ORGANISM: mouse

US-09-514-247A-8

Query Match 17.1%; Score 71.5; DB 4; Length 2441;

Best Local Similarity 36.5%; Pred. No. 20;

Matches 23; Conservative 4; Mismatches 33; Indels 3; Gaps 2;

Qy 4 SLLPRPAVPVPLRMQGPAPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVPSVSGRG 63

Db 843 SOLPCPPVTQSLHPTPPASTAAGM--PSLQHPATPGMTPPQPAATQ-PSTPVSSGQT 899

Qy 64 PDP 66

Db 900 PTP 902

RESULT 20

US-09-686-316-2

Sequence 2, Application US/09686316

Patent No. 6646115

GENERAL INFORMATION:

APPLICANT: Montminy, Marc R.

TITLE OF INVENTION: Methods for Treating Diabetes Mellitus

FILE REFERENCE: SALK1650-1

CURRENT APPLICATION NUMBER: US/09/686,316

CURRENT FILING DATE: 2000-10-10

PRIOR APPLICATION NUMBER: US/08/961,739

PRIOR FILING DATE: 1997-10-31

PRIOR APPLICATION NUMBER: US 194,468

PRIOR FILING DATE: 1994-02-10

NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2441
; TYPE: PRT
; ORGANISM: Mus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2441)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-686-316-2

Query Match 17.1%; Score 71.5; DB 4; Length 2441;
Best Local Similarity 36.5%; Pred. No. 20;
Matches 23; Conservative 4; Mismatches 33; Indels 3; Gaps 2;

QY 4 SLLPRAAVPVPLRMQGPAPVLSFLRPSWDLVSAFYSLPLAPLSPTSVISPVSVGRG 63
DB 843 SOLPCFVTQSLHPTPPASTAGM--PSLQHTAPQMTPOPAAPTQ-ESTVSSGQT 899
QY 64 PDP 66
DB 900 PTP 902

RESULT 21

US-09-252-991A-30479
; Sequence 30479, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30479
; LENGTH: 803
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30479

Query Match 17.0%; Score 71; DB 4; Length 803;
Best Local Similarity 36.4%; Pred. No. 5.6;
Matches 20; Conservative 8; Mismatches 17; Indels 10; Gaps 3;

QY 11 AVPVPLRMQGPAPVLSFLRPSWDLVSAFYSLPLAPLSPTSVISPVSVGRGP 64
DB 291 AVPVGARLQ----HPIGMA-----DLIGALHLLRLPDARPPVQVLFYVGQGF 336

RESULT 22

US-09-230-652-2
; Sequence 2, Application US/09230652A
; Patent No. 6537775
; GENERAL INFORMATION:
; APPLICANT: Tournier-Lasserre, Elisabeth
; APPLICANT: Joutel, Anne
; APPLICANT: Bousser, Marie-Germaine
; APPLICANT: Bach, Jean-Francois
; TITLE OF INVENTION: GENE INVOLVED IN CADASIL, METHOD OF DIAGNOSIS AND
; FILE REFERENCE: 03715.0048-00000
; CURRENT APPLICATION NUMBER: US/09/230,652A
; CURRENT FILING DATE: 1999-05-17
; EARLIER APPLICATION NUMBER: FR 96 09733
; EARLIER FILING DATE: 1996-08-01
; EARLIER APPLICATION NUMBER: FR 97 04680
; EARLIER FILING DATE: 1997-04-16

; EARLIER APPLICATION NUMBER: PCT/FR97/01433
; EARLIER FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2321
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human ADNC No. 6537775ch 3
US-09-230-652-2

Query Match 16.7%; Score 70; DB 4; Length 2321;
Best Local Similarity 37.7%; Pred. No. 28;
Matches 23; Conservative 2; Mismatches 24; Indels 12; Gaps 1;

QY 6 LRPRAAVPVPLRMQGPAPVLSFLRPSWDLVSAFYSLPLAPLSPTSVISPVSVGRGP 65
DB 2162 LLNPVAVPLDWARLPPAPPGPSFL-----LPLAPGOLLNFGTPVSPQERPP 2209
QY 66 P 66
DB 2210 P 2210

RESULT 23

US-09-147-236-11
; Sequence 11, Application US/09147236A
; Patent No. 6316251
; GENERAL INFORMATION:
; APPLICANT: TONOUCHI, Naoto
; APPLICANT: TSUCHIDA, Takayasu
; APPLICANT: YOSHINAGA, Fumihiro
; APPLICANT: TAHARA, Naoki
; APPLICANT: HAYASHI, Takahisa
; TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE
; FILE REFERENCE: 5537-011-0PCT
; CURRENT APPLICATION NUMBER: US/09/147,236A
; CURRENT FILING DATE: 1999-04-08
; EARLIER APPLICATION NUMBER: PCT/JP97/03633
; EARLIER FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Acetobacter xylinum
; FEATURE:
; OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, c, or
; OTHER INFORMATION: t
US-09-147-236-11

Query Match 16.6%; Score 69.5; DB 4; Length 344;
Best Local Similarity 31.0%; Pred. No. 2.9;
Matches 22; Conservative 7; Mismatches 19; Indels 23; Gaps 3;

QY 3 HSLPRAAVPVPLRMQGPAPVLSFL-----RPSWDLVSAFYSLPLAPLSPTSVPI 55
DB 158 YAAAPQPVATEVP--POPAPVAPVVAAPVQVRQERPS-----LSPVTPPK 201
QY 56 SPVSVGRGPD 66
DB 202 PAVSSFMARPP 212

RESULT 24

US-09-522-474-11
; Sequence 11, Application US/09522474
; Patent No. 6573076
; GENERAL INFORMATION:
; APPLICANT: TONOUCHI, Naoto
; APPLICANT: TSUCHIDA, Takayasu
; APPLICANT: YOSHINAGA, Fumihiro

APPLICANT: TAHARA, Naoki
APPLICANT: HAYASHI, Takahisa
TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE
FILE REFERENCE: 6537-011-OPCT
CURRENT APPLICATION NUMBER: US/09/522,474
CURRENT FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: US/09/147,236
PRIOR FILING DATE: 1999-04-08
PRIOR APPLICATION NUMBER: PCT/JP97/03633
PRIOR FILING DATE: 1997-10-09
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 344
TYPE: PRT
ORGANISM: Acetobacter xylinum
FEATURE:
OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, c, or
US-09-522-474-11

Query Match 16.6%; Score 69.5; DB 4; Length 344;
Best Local Similarity 31.0%; Pred. No. 2.9;
Matches 22; Conservative 7; Mismatches 19; Indels 23; Gaps 3;

QY 3 HSLPRPAAVPLRMQGPAPVLSFL-----RPSMDLVSFAFYSLLAPLSPTSVP 55
DB 158 YAAAPQVATVP--PQAPVAVVAQAQVRQRP-----LSFVTPK 201
QY 56 SPVSVGRGPD 66
DB 202 PAVSSEMAPRP 212

RESULT 25
US-08-615-170-21
Sequence 21, Application US/08615170
Patent No. 5776776
GENERAL INFORMATION:
APPLICANT: ORDAHL, Charles P.
APPLICANT: AZAKIE, Anthony
APPLICANT: MAR, Janet H.
APPLICANT: FARRANCE, Iain K.G.
APPLICANT: HALL, Deborah E.
APPLICANT: STEWART, Alexandre F.R.
APPLICANT: LARKIN, Sarah B.
TITLE OF INVENTION: DTEF-1 ISOFORMS AND USES THEREOF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/615,170
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01526
FILING DATE: 06-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/191,493
FILING DATE: 04-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 2307U-053120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid

NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 2307U-053120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-615-170-21

Query Match 16.6%; Score 69.5; DB 1; Length 432;
Best Local Similarity 35.0%; Pred. No. 3.8;
Matches 21; Conservative 11; Mismatches 17; Indels 11; Gaps 4;

QY 4 SLLPRPAAVPLRMQGPAPVLSFLRPSMDLVS-----AFYSLPLAPLSP--TSVPI 55
DB 155 SAAPRFWSGPIP--GQGPSQDIKPFAPQAPYIQPPMPPLASIE-PLAPLPAAASAVPV 211

RESULT 26
US-08-615-170-19
Sequence 19, Application US/08615170
Patent No. 5776776
GENERAL INFORMATION:
APPLICANT: ORDAHL, Charles P.
APPLICANT: AZAKIE, Anthony
APPLICANT: MAR, Janet H.
APPLICANT: FARRANCE, Iain K.G.
APPLICANT: HALL, Deborah E.
APPLICANT: STEWART, Alexandre F.R.
APPLICANT: LARKIN, Sarah B.
TITLE OF INVENTION: DTEF-1 ISOFORMS AND USES THEREOF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/615,170
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01526
FILING DATE: 06-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/191,493
FILING DATE: 04-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 2307U-053120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-615-170-19

Query Match      15.6%; Score 69.5; DB 1; Length 433;
Best Local Similarity 35.0%; Pred. No. 3.8; 17; Indels 11; Gaps 4;
Matches 21; Conservative 11; Mismatches 17; Indels 11; Gaps 4;

QY 4 SLPLPRPAAVPFLRMQPGPAHVPFLRPSWDLVS-----AFYSLPLAPLSP--TSVPI 55
Db 156 SAAPRFWSGRIIP--GQFGPSQDIKPPAQPAYPIQPPMPWPSLASVE-PLADLPAAASAVPV 212

RESULT 27
US-09-461-697-58
; Sequence 58, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Purnam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461.697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-697-58

Query Match      16.4%; Score 68.5; DB 3; Length 115;
Best Local Similarity 27.8%; Pred. No. 0.95; 29; Indels 27; Gaps 4;
Matches 25; Conservative 9; Mismatches 29; Indels 27; Gaps 4;

QY 1 GTHSLLPRAAIVPFLRMQP-----GPAH--PVLSEFLRPSWDL-----VSAFY 41
Db 11 GCHCWSRRDLTFLPHPSPEPGVLDCLGCPCHLLPLLSPGSPCWVLGLHFLSLHPPSAASASH 70

QY 42 SLPLAPLSPTSVPISPVS-----VGRG 63
Db 71 ALTITSLPPGLLPFGVVELTAHPQALIGRG 100

RESULT 28
US-09-461-697-54
; Sequence 54, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Purnam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461.697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0

US-09-461-697-54

Query Match      16.4%; Score 68.5; DB 3; Length 132;
Best Local Similarity 27.8%; Pred. No. 1.1; 29; Indels 27; Gaps 4;
Matches 25; Conservative 9; Mismatches 29; Indels 27; Gaps 4;

QY 1 GTHSLLPRAAIVPFLRMQP-----GPAH--PVLSEFLRPSWDL-----VSAFY 41
Db 28 GCHCWSRRDLTFLPHPSPEPGVLDCLGCPCHLLPLLSPGSPCWVLGLHFLSLHPPSAASASH 87

RESULT 29
US-09-461-697-52
; Sequence 52, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Purnam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461.697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-697-52

Query Match      16.4%; Score 68.5; DB 3; Length 139;
Best Local Similarity 27.8%; Pred. No. 1.2; 29; Indels 27; Gaps 4;
Matches 25; Conservative 9; Mismatches 29; Indels 27; Gaps 4;

QY 1 GTHSLLPRAAIVPFLRMQP-----GPAH--PVLSEFLRPSWDL-----VSAFY 41
Db 35 GCHCWSRRDLTFLPHPSPEPGVLDCLGCPCHLLPLLSPGSPCWVLGLHFLSLHPPSAASASH 94

QY 42 SLPLAPLSPTSVPISPVS-----VGRG 63
Db 95 ALTITSLPPGLLPFGVVELTAHPQALIGRG 124

RESULT 30
US-09-461-697-48
; Sequence 48, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Purnam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
```

```
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 456
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-697-48

Query Match      16.4%; Score 68.5; DB 3; Length 159;
Best Local Similarity 27.8%; Pred. No. 1.4;
Matches 25; Conservative 9; Mismatches 29; Indels 27; Gaps 4;

QY      1  CCHSLIPRPAAYVPLRWQP-----GPAH--PVLSLELPSPWDL-----VSAPY 41
Db      55  GCHCWSRRDLTELEHPSEPGVDCIGPCHLLPLSPGSCWVLGLHPSLHPPSAASASH 114

QY      42  SLPLAPLSPTSVFISPV-----VGRG 63
Db      115  ALTITSLPGLLPFVGVELTAHPQALIGRG 144

Search completed: July 4, 2004, 04:22:13
Job time : 6.23494 secs
```

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2004, 04:21:44 ; Search time 16.0221 Seconds
 (without alignments)
 1534.845 Million cell updates/sec

Title: US-09-506-079H-11
 Perfect score: 418
 Sequence: 1 GTHSLPRPAVPVPLRMP.....VGRGPDPAHVAVNLSRYEG 79

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 31128316 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 90 summaries

Database :	Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*	
2: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB.pep.*	
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*	
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*	
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*	
6: /cgn2_6/ptodata/2/pubpaa/PTU5_PUBCOMB.pep.*	
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*	
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*	
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*	
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*	
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*	
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*	
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*	
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*	
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*	
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*	
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*	
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	355	84.9	79	12	US-10-344-470-1
2	355	84.9	79	16	US-10-302-663-1
3	351	84.0	419	12	US-10-344-470-2
4	351	84.0	419	16	US-10-302-663-2
5	88	21.1	365	16	US-10-437-963-116009
6	86.5	20.7	1210	16	US-10-437-963-160288
7	86	20.6	459	10	US-09-468-147-206
8	86	20.6	459	10	US-09-468-147-207
9	86	20.6	459	12	US-10-319-745-206
10	86	20.6	459	12	US-10-319-745-207
11	85	20.3	427	16	US-10-437-963-203126
12	84	20.1	345	16	US-10-437-963-204246
13	83.5	20.0	888	10	US-09-931-836-35
14	83.5	20.0	888	12	US-10-147-493-544
15	83.5	20.0	888	12	US-10-145-127-544

16	83.5	20.0	888	12	US-10-160-503-544	Sequence 544, App
17	83.5	20.0	888	12	US-10-211-462-167	Sequence 167, App
18	83.5	20.0	888	12	US-10-143-118-544	Sequence 544, App
19	83.5	20.0	888	12	US-10-144-993-544	Sequence 544, App
20	83.5	20.0	888	12	US-10-158-787-544	Sequence 544, App
21	83.5	20.0	888	12	US-10-081-056-268	Sequence 268, App
22	83.5	20.0	888	12	US-10-140-024-544	Sequence 544, App
23	83.5	20.0	888	12	US-10-140-808-544	Sequence 544, App
24	83.5	20.0	888	12	US-10-152-405-544	Sequence 544, App
25	83.5	20.0	888	12	US-10-127-852A-544	Sequence 544, App
26	83.5	20.0	888	12	US-10-127-900A-544	Sequence 544, App
27	83.5	20.0	888	12	US-10-128-685A-544	Sequence 544, App
28	83.5	20.0	888	12	US-10-131-820A-544	Sequence 544, App
29	83.5	20.0	888	12	US-10-142-886-544	Sequence 544, App
30	83.5	20.0	888	12	US-10-146-728-544	Sequence 544, App
31	83.5	20.0	888	12	US-10-146-786-544	Sequence 544, App
32	83.5	20.0	888	12	US-10-147-499-544	Sequence 544, App
33	83.5	20.0	888	12	US-10-157-798-544	Sequence 544, App
34	83.5	20.0	888	12	US-10-305-654-268	Sequence 268, App
35	83.5	20.0	888	13	US-10-036-342-35	Sequence 35, Appl
36	83.5	20.0	888	13	US-10-036-041-35	Sequence 35, Appl
37	83.5	20.0	888	14	US-10-028-072-544	Sequence 544, App
38	83.5	20.0	888	14	US-10-035-855-35	Sequence 35, Appl
39	83.5	20.0	888	14	US-10-121-049-544	Sequence 544, App
40	83.5	20.0	888	14	US-10-123-904-544	Sequence 544, App
41	83.5	20.0	888	14	US-10-140-470-544	Sequence 544, App
42	83.5	20.0	888	14	US-10-175-746-544	Sequence 544, App
43	83.5	20.0	888	14	US-10-176-918-544	Sequence 544, App
44	83.5	20.0	888	14	US-10-176-921-544	Sequence 544, App
45	83.5	20.0	888	14	US-10-036-214-35	Sequence 35, Appl
46	83.5	20.0	888	14	US-10-137-865-544	Sequence 544, App
47	83.5	20.0	888	14	US-10-140-474-544	Sequence 544, App
48	83.5	20.0	888	14	US-10-035-719-35	Sequence 35, Appl
49	83.5	20.0	888	14	US-10-142-431-544	Sequence 544, App
50	83.5	20.0	888	14	US-10-143-114-544	Sequence 544, App
51	83.5	20.0	888	14	US-10-140-002-544	Sequence 544, App
52	83.5	20.0	888	14	US-10-036-160-35	Sequence 35, Appl
53	83.5	20.0	888	14	US-10-142-419-544	Sequence 544, App
54	83.5	20.0	888	14	US-10-035-958-35	Sequence 35, Appl
55	83.5	20.0	888	14	US-10-036-150-35	Sequence 35, Appl
56	83.5	20.0	888	14	US-10-123-262-544	Sequence 544, App
57	83.5	20.0	888	14	US-10-142-423-544	Sequence 544, App
58	83.5	20.0	888	14	US-10-121-050-544	Sequence 544, App
59	83.5	20.0	888	14	US-10-141-755-544	Sequence 544, App
60	83.5	20.0	888	14	US-10-143-032-544	Sequence 544, App
61	83.5	20.0	888	14	US-10-123-108-544	Sequence 544, App
62	83.5	20.0	888	14	US-10-123-236-544	Sequence 544, App
63	83.5	20.0	888	14	US-10-123-261-544	Sequence 544, App
64	83.5	20.0	888	14	US-10-140-921-544	Sequence 544, App
65	83.5	20.0	888	14	US-10-140-928-544	Sequence 544, App
66	83.5	20.0	888	14	US-10-121-045-544	Sequence 544, App
67	83.5	20.0	888	14	US-10-123-292-544	Sequence 544, App
68	83.5	20.0	888	14	US-10-123-903-544	Sequence 544, App
69	83.5	20.0	888	14	US-10-124-819-544	Sequence 544, App
70	83.5	20.0	888	14	US-10-124-822-544	Sequence 544, App
71	83.5	20.0	888	14	US-10-140-925-544	Sequence 544, App
72	83.5	20.0	888	14	US-10-160-498-544	Sequence 544, App
73	83.5	20.0	888	14	US-10-124-824-544	Sequence 544, App
74	83.5	20.0	888	14	US-10-127-825A-544	Sequence 544, App
75	83.5	20.0	888	14	US-10-127-829A-544	Sequence 544, App
76	83.5	20.0	888	14	US-10-127-835A-544	Sequence 544, App
77	83.5	20.0	888	14	US-10-127-839A-544	Sequence 544, App
78	83.5	20.0	888	14	US-10-127-901A-544	Sequence 544, App
79	83.5	20.0	888	14	US-10-128-693A-544	Sequence 544, App
80	83.5	20.0	888	14	US-10-131-813A-544	Sequence 544, App
81	83.5	20.0	888	14	US-10-131-818A-544	Sequence 544, App
82	83.5	20.0	888	14	US-10-131-823A-544	Sequence 544, App
83	83.5	20.0	888	14	US-10-131-824A-544	Sequence 544, App
84	83.5	20.0	888	14	US-10-131-830A-544	Sequence 544, App
85	83.5	20.0	888	14	US-10-131-837A-544	Sequence 544, App
86	83.5	20.0	888	14	US-10-137-872A-544	Sequence 544, App
87	83.5	20.0	888	14	US-10-147-500-544	Sequence 544, App
88	83.5	20.0	888	14	US-10-147-502-544	Sequence 544, App

89 83.5 20.0 888 14 US-10-147-515-544 Sequence 544, App
90 83.5 20.0 888 14 US-10-147-517-544 Sequence 544, App

ALIGNMENTS

RESULT 1
US-10-344-470-1
; Sequence 1, Application US/10344470
; Publication No. US20040052796A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; TITLE OF INVENTION: EXPRESSION OF HERSTATIN, AN ALTERNATIVE TO HER-2/NEU PRODUCT, IN C
; TITLE OF INVENTION: THAT EXPRESS EITHER P185HER-2 OR THE EGF RECEPTOR INHIBITS RECH
; TITLE OF INVENTION: CELL GROWTH
; FILE REFERENCE: 49321-81
; CURRENT APPLICATION NUMBER: US/10/344,470
; CURRENT FILING DATE: 2003-06-09
; PRIOR APPLICATION NUMBER: PCT / US01/25502
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 09/638,834
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC FEATURE
; LOCATION: (2)..(2)
; OTHER INFORMATION: Applicants herein disclose Thr and Ser sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (5)..(5)
; OTHER INFORMATION: Applicants herein disclose Leu and Pro sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (6)..(6)
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (16)..(16)
; OTHER INFORMATION: Applicants herein disclose Leu and Gln sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (18)..(18)
; OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (21)..(21)
; OTHER INFORMATION: Applicants herein disclose Gly, Asp, Ala and Val sequence variant
; OTHER INFORMATION: s at this position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (36)..(36)
; OTHER INFORMATION: Applicants herein disclose Leu and Ile sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (54)..(54)
; OTHER INFORMATION: Applicants herein disclose Pro and Arg sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (64)..(64)
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this

; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (73)..(73)
; OTHER INFORMATION: Applicants herein disclose Asp and Asn sequence variants at thi
; OTHER INFORMATION: position
US-10-344-470-1

Query Match 84.9%; Score 355; DB 12; Length 79;
Best Local Similarity 87.3%; Pred. No. 1e-27;
Matches 59; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GTHSLPRDAVPVLRQPGAHVPLSFLRPSWDLVSAFYSLPLAPLSPTSVPIPSV 60
DB 1 CXHXXPRDAVPVLRQPGAHVPLSFLRPSWDLVSAFYSLPLAPLSPTSVPIPSV 60

QY 61 GRGPDPAHVAVNLSRYEG 79
DB 61 GRGPDPAHVAVNLSRYEG 79

RESULT 2
US-10-302-663-1
; Sequence 1, Application US/10302663
; Publication No. US20040022785A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; TITLE OF INVENTION: EXPRESSION OF HERSTATIN, AN ALTERNATIVE HER-2/NEU PRODUCT, IN C
; TITLE OF INVENTION: EXPRESS EITHER P185HER-2 OR THE EGF RECEPTOR INHIBITS RECEPTOR
; TITLE OF INVENTION: GROWTH
; FILE REFERENCE: 49321-73
; CURRENT APPLICATION NUMBER: US/10/302,663
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 09/638,834
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 1
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 2
; OTHER INFORMATION: Applicants herein disclose Thr and Ser sequence variants at thi
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 5
; OTHER INFORMATION: Applicants herein disclose Leu and Pro sequence variants at thi
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 6
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at thi
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 16
; OTHER INFORMATION: Applicants herein disclose Leu and Gln sequence variants at thi
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 18
; OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at thi
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 21
; OTHER INFORMATION: Applicants herein disclose Gly, Asp, Ala and Val sequence varia
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 36
; OTHER INFORMATION: Applicants herein disclose Leu and Ile sequence variants at thi
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 54
; OTHER INFORMATION: Applicants herein disclose Pro and Arg sequence variants at thi


```

; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 64
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 73
; OTHER INFORMATION: Applicants herein disclose Asp and Asn sequence variants at this
US-10-302-663-1

Query Match      84.9%; Score 355; DB 16; Length 79;
Best Local Similarity 87.3%; Pred. No. 1e-27;
Matches 69; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GTGSLLRPAAPVPLRQPCPAHPVLSTFLRPSNDLYSAFYSLPLAPLSPTSVDISPSV 60
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GXSXXPRPAAPVPPVPEXQPAHPVLSTFLRPSNDXVSAFYSLPLAPLSPTSVXISPSV 60
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 GRGPDPPDAHVAVNLSRYEG 79
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GRGPDPPDAHVAVNLSRYEG 79
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
US-10-344-470-2
; Sequence 2, Application US/10344470
; Publication No. US20040052796A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; TITLE OF INVENTION: EXPRESSION OF HERSTATIN, AN ALTERNATIVE TO HER-2/NEU PRODUCT, IN
; TITLE OF INVENTION: THAT EXPRESS EITHER p185HER-2 OR THE EGF RECEPTOR INHIBITS RECEL
; TITLE OF INVENTION: CELL GROWTH
; FILE REFERENCE: 49321-81
; CURRENT APPLICATION NUMBER: US/10/344,470
; CURRENT FILING DATE: 2003-06-09
; PRIOR APPLICATION NUMBER: PCT / US01/25502
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 09/638,834
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (342)..(342)
; OTHER INFORMATION: Applicants herein disclose Thr and Ser sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (345)..(345)
; OTHER INFORMATION: Applicants herein disclose Leu and Pro sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (346)..(346)
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (356)..(356)
; OTHER INFORMATION: Applicants herein disclose Leu and Gln sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (358)..(358)
; OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (361)..(361)

```

```

; NAME/KEY: VARIANT
; LOCATION: 358
; OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 361
; OTHER INFORMATION: Applicants herein disclose Gly, Asp, Ala and Val sequence variant
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 376
; OTHER INFORMATION: Applicants herein disclose Leu and Ile sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 394
; OTHER INFORMATION: Applicants herein disclose Pro and Arg sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 404
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 413
; OTHER INFORMATION: Applicants herein disclose Asp and Asn sequence variants at this
; OTHER INFORMATION:
US-10-302-663-2

Query Match      84.0%; Score 351; DB 16; Length 419;
Best Local Similarity 86.1%; Pred. No. 1.5e-26;
Matches 68; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 GTHSLPRPAVPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSVPIPSVSV 60
DB 341 GXHXXPRPAVPVFXRXQPAHPVLSFLRPSWDXVSAFYSLPLAPLDTSVVAISPSV 400

QY 61 GRGPDPAHVAVNLRSYEG 79
DB 401 GRGXDPDAHVAVXLSRYEG 419

RESULT 5
US-10-437-963-116009
; Sequence 116009, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 116009
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_19551C.1.pep
US-10-437-963-116009

Query Match      21.1%; Score 88; DB 16; Length 365;
Best Local Similarity 28.7%; Pred. No. 1.2;
Matches 29; Conservative 4; Mismatches 36; Indels 32; Gaps 2;

QY 3 HSLJPRPAVPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSVPIPSVSV 50
DB 102 HPRVPRAPAPVPPAPPTPTPTPALAPPADVPFGFMPPPLVPTTPPAHAPPVATPPTL 161
```

```

QY 51 -TSVPISPVSV-----VGRGPDPAHVA 71
DB 162 ATVPVTPPTATPADVPVPSAMPAAAAAPLVLRGPAHRLVS 202

RESULT 6
US-10-437-963-160288
; Sequence 160288, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 160288
; LENGTH: 1410
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1410)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_59580C.1.pep
US-10-437-963-160288

Query Match      20.7%; Score 86.5; DB 16; Length 1410;
Best Local Similarity 33.3%; Pred. No. 7.3;
Matches 25; Conservative 7; Mismatches 28; Indels 15; Gaps 2;

QY 4 SLLPRPAVPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSVPIPSVSVG-R 62
DB 1124 SLPPPPPPPLPSGPPQPAPPPLPIQP-----PIPPPPVPSFSLGYQ 1169

QY 63 GPDPAHVAVNLRSY 77
DB 1170 PPAPEYFRASNIQY 1184

RESULT 7
US-09-468-147-206
; Sequence 206, Application US/09468147A
; Publication No. US20030049601A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; FILE REFERENCE: 6232.US.P1
; CURRENT APPLICATION NUMBER: US/09/468,147A
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/173,141
; EARLIER FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: US 60/061,199
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 206
```

```
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: CKSORF32M-3.pep
US-09-468-147-206

Query Match      20.6%; Score 86; DB 10; Length 459;
Best Local Similarity 34.2%; Pred. No. 2.5;
Matches 27; Conservative 12; Mismatches 28; Indels 12; Gaps 4;

Qy 1 GTHSLPRPAAVPELRMQGPAHPVLSFLRPSWDLVSAFYSLPLAPL---SPTSVPISP 57
Db 57 GVTGLILSPS--PSPIFIQTPFS-PPMSFHNFGLEALDSRPAPLAPLGVTSPPSAPPLPP 113

Qy 58 V-----SVGRGPDPAHV 70
Db 114 VVDLPQLGLRGADGTAEL 132

RESULT 8
US-09-468-147-207
; Sequence 207, Application US/09468147A
; Publication No. US20030049601A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232 US.P1
; CURRENT APPLICATION NUMBER: US/09/468,147A
; CURRENT FILING DATE: 1999-12-21
; EARLIER FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: US 60/061,199
; EARLIER FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 207
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: FLORF32M-14-5.ppe
US-09-468-147-207

Query Match      20.6%; Score 86; DB 10; Length 459;
Best Local Similarity 34.2%; Pred. No. 2.5;
Matches 27; Conservative 12; Mismatches 28; Indels 12; Gaps 4;

Qy 1 GTHSLPRPAAVPELRMQGPAHPVLSFLRPSWDLVSAFYSLPLAPL---SPTSVPISP 57
Db 57 GVTGLILSPS--PSPIFIQTPFS-PPMSFHNFGLEALDSRPAPLAPLGVTSPPSAPPLPP 113

Qy 58 V-----SVGRGPDPAHV 70
Db 114 VVDLPQLGLRGADGTAEL 132

RESULT 9
US-10-319-745-206
; Sequence 206, Application US/10319745
; Publication No. US20030211467A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232 US.P1
; CURRENT APPLICATION NUMBER: US/10/319,745
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: US/09/468,147A
; PRIOR FILING DATE: 1999-12-21
; PRIOR FILING DATE: 1998-10-15
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-10-15
; PRIOR FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 207
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: FLORF32M-14-5.ppe
US-10-319-745-207

Query Match      20.6%; Score 86; DB 12; Length 459;
Best Local Similarity 34.2%; Pred. No. 2.5;
Matches 27; Conservative 12; Mismatches 28; Indels 12; Gaps 4;

Qy 1 GTHSLPRPAAVPELRMQGPAHPVLSFLRPSWDLVSAFYSLPLAPL---SPTSVPISP 57
Db 57 GVTGLILSPS--PSPIFIQTPFS-PPMSFHNFGLEALDSRPAPLAPLGVTSPPSAPPLPP 113

Qy 58 V-----SVGRGPDPAHV 70
Db 114 VVDLPQLGLRGADGTAEL 132

RESULT 10
US-10-319-745-207
; Sequence 207, Application US/10319745
; Publication No. US20030211467A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232 US.P1
; CURRENT APPLICATION NUMBER: US/10/319,745
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: US/09/468,147A
; PRIOR FILING DATE: 1999-12-21
; PRIOR FILING DATE: 1998-10-15
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-10-15
; PRIOR FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 207
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: FLORF32M-14-5.ppe
US-10-319-745-207

Query Match      20.6%; Score 86; DB 12; Length 459;
Best Local Similarity 34.2%; Pred. No. 2.5;
Matches 27; Conservative 12; Mismatches 28; Indels 12; Gaps 4;

Qy 1 GTHSLPRPAAVPELRMQGPAHPVLSFLRPSWDLVSAFYSLPLAPL---SPTSVPISP 57
Db 57 GVTGLILSPS--PSPIFIQTPFS-PPMSFHNFGLEALDSRPAPLAPLGVTSPPSAPPLPP 113

Qy 58 V-----SVGRGPDPAHV 70
Db 114 VVDLPQLGLRGADGTAEL 132

RESULT 11
US-10-319-745-207
; Sequence 207, Application US/10319745
; Publication No. US20030211467A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Schlauder, George G.
; APPLICANT: Erker, James C.
; APPLICANT: Desai, Suresh M.
; APPLICANT: Dawson, George J.
; APPLICANT: Mushahwar, I. K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: HEPATITIS E VIRUS
; FILE REFERENCE: 6232 US.P1
; CURRENT APPLICATION NUMBER: US/10/319,745
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: US/09/468,147A
; PRIOR FILING DATE: 1999-12-21
; PRIOR FILING DATE: 1998-10-15
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-10-15
; PRIOR FILING DATE: 1997-10-15
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 207
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Hepatitis E Virus
; FEATURE:
; OTHER INFORMATION: CKSORF32M-3.ppe
US-10-319-745-206

Query Match      20.6%; Score 86; DB 12; Length 459;
Best Local Similarity 34.2%; Pred. No. 2.5;
Matches 27; Conservative 12; Mismatches 28; Indels 12; Gaps 4;

Qy 1 GTHSLPRPAAVPELRMQGPAHPVLSFLRPSWDLVSAFYSLPLAPL---SPTSVPISP 57
Db 57 GVTGLILSPS--PSPIFIQTPFS-PPMSFHNFGLEALDSRPAPLAPLGVTSPPSAPPLPP 113

Qy 58 V-----SVGRGPDPAHV 70
Db 114 VVDLPQLGLRGADGTAEL 132
```

Db 57 GVTGLILSPS--PSPIFIQTPS-PPMSFFHNGLELALDSRPAPLAPLGVTSPSAPPLPP 113

QY 58 V-----SVGRGPPDDAHV 70

Db 114 VVDLPQZGLRRGADGTAEI 132

RESULT 11

US-10-437-963-203126

Sequence 203126, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 203126

LENGTH: 427

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(427)

OTHER INFORMATION: unsure at all Xaa locations

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_9833C.1.pap

US-10-437-963-203126

Query Match 20.3%; Score 85; DB 16; Length 427;

Best Local Similarity 35.1%; Pred. No. 2.9;

Mismatches 26; Conservative 12; Mismatches 24; Indels 12; Gaps 4;

QY 9 PAAVPLRMQGFPAHVLSTLRPS---WDLVSIFY---SLPLAPLSPTSPVSPVSGR 62

Db 20 PAAIPVAV-----PSPVLGSL-PSAINKWDLTPFYRNPIFSLPVVPSPLPALLSLPR 73

QY 63 GPDPDAHVAVNLNR 76

Db 74 RRPAPACAASTVR 87

RESULT 12

US-10-437-963-204246

Sequence 204246, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 204246

LENGTH: 345

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(345)

OTHER INFORMATION: unsure at all Xaa locations

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_99352C.1.pap

US-10-437-963-204246

Query Match 20.1%; Score 84; DB 16; Length 345;

Best Local Similarity 31.1%; Pred. No. 2.9;

Mismatches 23; Conservative 10; Mismatches 23; Indels 18; Gaps 3;

QY 8 RPAAVPLRMQGFPAHVLSTLRP-----SWDLVSIFYSLPLAPLSPTSPVSPVSV--- 60

Db 57 RPPSLSTP---HPSPVPVLLKPKRPSPPSPAGHHRRPPPPINPSALPLPTFVHRP 113

QY 61 -----GRGDDP 66

Db 114 LPTPLRWGRPPDP 127

RESULT 13

US-09-931-836-35

Sequence 35, Application US/09931836

Publication No. US20030027249A1

GENERAL INFORMATION:

APPLICANT: Desnoyers, Luc

APPLICANT: Baton, Dan L.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Stewart, Timothy A.

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3030R1C1

CURRENT APPLICATION NUMBER: US/09/931,836

CURRENT FILING DATE: 2001-08-16

PRIOR APPLICATION NUMBER: 60/085579

PRIOR FILING DATE: 1998-05-15

PRIOR APPLICATION NUMBER: 60/112514

PRIOR FILING DATE: 1998-12-15

PRIOR APPLICATION NUMBER: 60/113300

PRIOR FILING DATE: 1998-12-22

PRIOR APPLICATION NUMBER: 60/113430

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: 60/113605

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: 60/113621

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: 60/114140

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: 60/115552

PRIOR FILING DATE: 1999-01-12

PRIOR APPLICATION NUMBER: 60/116843

PRIOR FILING DATE: 1999-01-22

PRIOR APPLICATION NUMBER: 60/125774

PRIOR FILING DATE: 1999-03-23

PRIOR APPLICATION NUMBER: 60/125778

PRIOR FILING DATE: 1999-03-23

PRIOR APPLICATION NUMBER: 60/125826

PRIOR FILING DATE: 1999-03-24

PRIOR APPLICATION NUMBER: 60/127035

PRIOR FILING DATE: 1999-03-31

PRIOR APPLICATION NUMBER: 60/127706

PRIOR FILING DATE: 1999-04-05

PRIOR APPLICATION NUMBER: 60/129122

PRIOR FILING DATE: 1999-04-13

PRIOR APPLICATION NUMBER: 60/130359

;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/131270
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131272
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131291
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/132371
;; PRIOR FILING DATE: 1999-05-04
;; PRIOR APPLICATION NUMBER: 60/132379
;; PRIOR FILING DATE: 1999-05-04
;; PRIOR APPLICATION NUMBER: 60/132383
;; PRIOR FILING DATE: 1999-05-04
;; PRIOR APPLICATION NUMBER: 60/135750
;; PRIOR FILING DATE: 1999-05-25
;; PRIOR APPLICATION NUMBER: 60/136166
;; PRIOR FILING DATE: 1999-06-08
;; PRIOR APPLICATION NUMBER: 60/144791
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 60/146970
;; PRIOR FILING DATE: 1999-08-03
;; PRIOR APPLICATION NUMBER: 60/162506
;; PRIOR FILING DATE: 1999-10-29
;; PRIOR APPLICATION NUMBER: 09/311832
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 09/380142
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/644848
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 09/747259
;; PRIOR FILING DATE: 2000-12-20
;; PRIOR APPLICATION NUMBER: 09/816744
;; PRIOR FILING DATE: 2001-03-22
;; PRIOR APPLICATION NUMBER: 09/854208
;; PRIOR FILING DATE: 2001-05-10
;; PRIOR APPLICATION NUMBER: 09/854280
;; PRIOR FILING DATE: 2001-05-10
;; PRIOR APPLICATION NUMBER: 09/874503
;; PRIOR FILING DATE: 2001-06-05
;; PRIOR APPLICATION NUMBER: 09/869599
;; PRIOR FILING DATE: 2001-06-29
;; PRIOR APPLICATION NUMBER: 09/908,827
;; PRIOR FILING DATE: 2001-07-18
;; PRIOR APPLICATION NUMBER: PCT/US99/10733
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: PCT/US99/28551
;; PRIOR FILING DATE: 1999-12-02
;; PRIOR APPLICATION NUMBER: PCT/US99/30720
;; PRIOR FILING DATE: 1999-12-22
;; PRIOR APPLICATION NUMBER: PCT/US00/05601
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: PCT/US00/05841
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: PCT/US00/14042
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: PCT/US00/15264
;; PRIOR FILING DATE: 2000-06-02
;; PRIOR APPLICATION NUMBER: PCT/US00/23522
;; PRIOR FILING DATE: 2000-08-23
;; PRIOR APPLICATION NUMBER: PCT/US00/23328
;; PRIOR FILING DATE: 2000-08-24
;; PRIOR APPLICATION NUMBER: PCT/US00/32678
;; PRIOR FILING DATE: 2000-12-01
;; PRIOR APPLICATION NUMBER: PCT/US00/34956
;; PRIOR FILING DATE: 2000-12-20
;; PRIOR APPLICATION NUMBER: PCT/US01/06520
;; PRIOR FILING DATE: 2001-08-28
;; PRIOR APPLICATION NUMBER: PCT/US01/17800
;; PRIOR FILING DATE: 2001-06-01
;; PRIOR APPLICATION NUMBER: PCT/US01/19692
;; PRIOR FILING DATE: 2001-06-20
;; PRIOR APPLICATION NUMBER: PCT/US01/21066
;; PRIOR FILING DATE: 2001-06-29

;; PRIOR APPLICATION NUMBER: PCT/US01/21735
;; PRIOR FILING DATE: 2001-07-09
;; NUMBER OF SEQ ID NOS: 80
;; SEQ ID NO 35
;; LENGTH: 888
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-09-931-836-375

Query Match 20.0%; Score 83.5; DB 10; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;
QY 1 GTH-----SILPRPAAPVPLRMQGPAPVLSFLRP-SWD-----LVSAFYSLPAPLSPT 51
Db 698 GPHDLDSGLLTPEQTPLPQKRLPTP-RPHPHALGPRANDHGHPLLPASASSLLLLAPA 756
QY 52 SVPISPVSVGRGPPDPAHV 70
Db 757 RAPEQPPAPGE-PTPDGRL 774

RESULT 14

US-10-147-493-544
;; Sequence 544, Application US/10147493
;; Publication No. US20040029217A1
;; GENERAL INFORMATION:

;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Beresini, Maureen
;; APPLICANT: DeForge, Laura
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Gao, Wei-Qiang
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Sherwood, Steven
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tamas, Daniel
;; APPLICANT: Watanabe, Colin K
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; TITLE OF INVENTION: ACIDS ENCODING THE SAME
;; FILE REFERENCE: P3330R1C345
;; CURRENT APPLICATION NUMBER: US/10/147,493
;; PRIOR APPLICATION NUMBER: 2002-05-17
;; PRIOR Application removed - See File Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 550
;; SEQ ID NO 544
;; LENGTH: 888
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-147-493-544

Query Match 20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;
QY 1 GTH-----SILPRPAAPVPLRMQGPAPVLSFLRP-SWD-----LVSAFYSLPAPLSPT 51
Db 698 GPHDLDSGLLTPEQTPLPQKRLPTP-RPHPHALGPRANDHGHPLLPASASSLLLLAPA 756
QY 52 SVPISPVSVGRGPPDPAHV 70
Db 757 RAPEQPPAPGE-PTPDGRL 774

RESULT 15

US-10-145-127-544
;; Sequence 544, Application US/10145127

```
; Publication No. US20040033558A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC252
; CURRENT APPLICATION NUMBER: US/10/145,127
; CURRENT FILING DATE: 2002-05-13
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-145-127-544

Query Match      20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

QY 1 GTH-----SLLPRPAAVVPVLRMQGPAHPVLSFLRP-SWD-----LVSAFYSLPLAPLSPT 51
Db 698 GPHDLDSGLLPTPEQTPLPQKRLPTP-HPHPHALGPRAWDHGHPLLPASASSLLLLAPA 756

QY 52 SVPISPVSVGREGDPDAHV 70
Db 757 RAPEQPPAPGE-PTPDGRL 774

RESULT 16
US-10-160-503-544
; Sequence 544, Application US/10/160503
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC446
; CURRENT APPLICATION NUMBER: US/10/160,503
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 550

Query Match      20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

QY 1 GTH-----SLLPRPAAVVPVLRMQGPAHPVLSFLRP-SWD-----LVSAFYSLPLAPLSPT 51
Db 698 GPHDLDSGLLPTPEQTPLPQKRLPTP-HPHPHALGPRAWDHGHPLLPASASSLLLLAPA 756

QY 52 SVPISPVSVGREGDPDAHV 70
Db 757 RAPEQPPAPGE-PTPDGRL 774

RESULT 17
US-10-211-462-167
; Sequence 167, Application US/10211462
; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
; FILE REFERENCES: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 167
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-462-167

Query Match      20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

QY 1 GTH-----SLLPRPAAVVPVLRMQGPAHPVLSFLRP-SWD-----LVSAFYSLPLAPLSPT 51
Db 698 GPHDLDSGLLPTPEQTPLPQKRLPTP-HPHPHALGPRAWDHGHPLLPASASSLLLLAPA 756

QY 52 SVPISPVSVGREGDPDAHV 70
Db 757 RAPEQPPAPGE-PTPDGRL 774

RESULT 18
US-10-143-118-544
; Sequence 544, Application US/10143118
; Publication No. US20040038335A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
```

```

; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC228
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-143-118-544

Query Match 20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

QY 1 GTH----SLLPRPAVPVPLRMQGPAPHPVLSFLRP-SWD----LVSAFYSLPLAPLSPT 51
DB 698 GPHDLDSGLLPTEQTPLPQKRLTP-HPHPALGPRAWDHGHPLLPASASSLLLLAPA 756
QY 52 SVPISPVSVGRGPDPAHV 70
DB 757 RAPEQPPAPGE-PTPDGRL 774
```

```

RESULT 19
US-10-144-993-544
; Sequence 544, Application US/10144993
; Publication No. US20040038336A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC261
; CURRENT APPLICATION NUMBER: US/10/144,993
; CURRENT FILING DATE: 2002-05-13
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-144-993-544
```

```

Query Match 20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

QY 1 GTH----SLLPRPAVPVPLRMQGPAPHPVLSFLRP-SWD----LVSAFYSLPLAPLSPT 51
DB 698 GPHDLDSGLLPTEQTPLPQKRLTP-HPHPALGPRAWDHGHPLLPASASSLLLLAPA 756
QY 52 SVPISPVSVGRGPDPAHV 70
DB 757 RAPEQPPAPGE-PTPDGRL 774

RESULT 20
US-10-158-787-544
; Sequence 544, Application US/10158787
; Publication No. US20040039164A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC449
; CURRENT APPLICATION NUMBER: US/10/158,787
; CURRENT FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-158-787-544

Query Match 20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

QY 1 GTH----SLLPRPAVPVPLRMQGPAPHPVLSFLRP-SWD----LVSAFYSLPLAPLSPT 51
```

Db 698 GPHDLSGLLPTFQTPFLKRLPTP-HPHPHALGPRANDHGHPLLPASASSSLLLLLAPA 756
Qy 52 SVPISPVSVGRGPDPAHV 70
Db 757 RAPEQPPAPGE-PTPDGRL 774

RESULT 21

US-10-081-056-268
; Sequence 268, Application US/10081056
; Publication No. US20040043927A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Ye, Weilan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS
; FILE REFERENCE: P3235P1C1
; CURRENT APPLICATION NUMBER: US/10/081.056
; CURRENT FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: PCT/US01/21735
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/219,556
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/220,624
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/220,664
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/222,695
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: PCT/US00/23522
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/230,978
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/000,000
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 09/664,610
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/242,922
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 09/709,238
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/30952
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/30873
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/747,259
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/34956

; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: US 09/767,609
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: US 09/796,498
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06666
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: US 09/802,706
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 09/808,689
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 09/816,744
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 09/828,366
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/854,208
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 09/854,280
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 09/866,028
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 09/866,034
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: PCT/US01/17092
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 09/870,574
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US01/17443
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: PCT/US01/17800
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US01/19692
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: PCT/US01/00000
; PRIOR FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 383
; SEQ ID NO 268
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homosapiens
US-10-081-056-268
Query Match 20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;
Qy 1 GTH-----SLLPRPAVVFELRMQGPAPHPVLSFLRP-SWD---LVSATYSLPLAPLSPT 51
Db 698 GPHDLSGLLPTFQTPFLKRLPTP-HPHPHALGPRANDHGHPLLPASASSSLLLLLAPA 756
Qy 52 SVPISPVSVGRGPDPAHV 70
Db 757 RAPEQPPAPGE-PTPDGRL 774
RESULT 22
US-10-140-024-544
; Sequence 544, Application US/10140024
; Publication No. US20040058424A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven

Qy 52 SVPISPVSVGRGPDPAHV 70
Db 757 RAPEQPPAPGE-PTPDGRL 774

RESULT 24

US-10-152-405-544
; Sequence 544, Application US/10152405
; Publication No. US20030211571A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC383
; CURRENT APPLICATION NUMBER: US/10/152,405
; CURRENT FILING DATE: 2002-05-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-152-405-544

Query Match 20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

Qy 1 GTH-----SLLPRPAAVPFLRMQGPAPHPVLSFLRP-SWD-----LVSAFYSLPLAPLSPT 51
Db 698 GPHDLSGLLPTPEQTLPQKRLPTP-HPHPHALGPRAWDHGHPLLPASASSLLLLAPA 756

Qy 52 SVPISPVSVGRGPDPAHV 70
Db 757 RAPEQPPAPGE-PTPDGRL 774

RESULT 25

US-10-127-852A-544
; Sequence 544, Application US/10127852A
; Publication No. US20030203428A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.

; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC69
; CURRENT APPLICATION NUMBER: US/10/140,024
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-024-544

Query Match 20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

Qy 1 GTH-----SLLPRPAAVPFLRMQGPAPHPVLSFLRP-SWD-----LVSAFYSLPLAPLSPT 51
Db 698 GPHDLSGLLPTPEQTLPQKRLPTP-HPHPHALGPRAWDHGHPLLPASASSLLLLAPA 756

Qy 52 SVPISPVSVGRGPDPAHV 70
Db 757 RAPEQPPAPGE-PTPDGRL 774

RESULT 23

US-10-140-808-544
; Sequence 544, Application US/10140808
; Publication No. US20030017563A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC182
; CURRENT APPLICATION NUMBER: US/10/140,808
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-808-544

Query Match 20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

Qy 1 GTH-----SLLPRPAAVPFLRMQGPAPHPVLSFLRP-SWD-----LVSAFYSLPLAPLSPT 51

RESULT 26

US-10-127-852A-544
; Sequence 544, Application US/10127852A
; Publication No. US20030203428A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.

```

; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C8
; CURRENT APPLICATION NUMBER: US/10/127,852A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-127-852A-544

Query Match 20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

QY 1 GTH----SLLPRPAAVPLRMQGPAPVLSFLRP-SWD---LVSAFYSLPLAPLSPT 51
Db 698 GPHLDGSLLPETPQTPLFQKRLPTP-HPHPHALGPRANDHGHPLLPASASSLLLLAPA 756

QY 52 SVPISPVSVGRGPDPAHV 70
Db 757 RAPEQPPARG-PTPDGRL 774

RESULT 26
US-10-127-900A-544
; Sequence 544, Application US/10127900A
; Publication No. US20030203429A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C116
; CURRENT APPLICATION NUMBER: US/10/128,685A
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: 60/049911

```

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C81
; CURRENT APPLICATION NUMBER: US/10/127,900A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-127-900A-544

Query Match 20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

QY 1 GTH----SLLPRPAAVPLRMQGPAPVLSFLRP-SWD---LVSAFYSLPLAPLSPT 51
Db 698 GPHLDGSLLPETPQTPLFQKRLPTP-HPHPHALGPRANDHGHPLLPASASSLLLLAPA 756

QY 52 SVPISPVSVGRGPDPAHV 70
Db 757 RAPEQPPARG-PTPDGRL 774

RESULT 27
US-10-128-685A-544
; Sequence 544, Application US/10128685A
; Publication No. US20030203430A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C116
; CURRENT APPLICATION NUMBER: US/10/128,685A
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: 60/049911

```

```
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-128-685A-544

Query Match      20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

QY 1 GTH----SLLPRPAAVPVLRMQCGPAHPVLSFLRP-SWD----LVSAFYSLPLAPLSPT 51
DB 698 GFHDLDSGLLTPTEQTPQLPKRLTP-RPHPHALGPRAWDHGHPLIPASASSLLLLAPA 756

QY 52 SVPIPSVSVGRGPDPAHV 70
DB 757 RAPEQPPAPGE-PTPDGRL 774

RESULT 28
US-10-131-820A-544
; Sequence 544, Application US/10131820A
; Publication No. US20030203431A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C144
; CURRENT APPLICATION NUMBER: US/10/131,820A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
```

```
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-820A-544

Query Match      20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

QY 1 GTH----SLLPRPAAVPVLRMQCGPAHPVLSFLRP-SWD----LVSAFYSLPLAPLSPT 51
DB 698 GFHDLDSGLLTPTEQTPQLPKRLTP-RPHPHALGPRAWDHGHPLIPASASSLLLLAPA 756

QY 52 SVPIPSVSVGRGPDPAHV 70
DB 757 RAPEQPPAPGE-PTPDGRL 774

RESULT 29
US-10-142-886-544
; Sequence 544, Application US/10142886
; Publication No. US20030203432A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C236
; CURRENT APPLICATION NUMBER: US/10/142,886
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-886-544

Query Match      20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. No. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;
```

Qy 1 GTH----SLPRPAAVPLRMQGPAPHVLSFLRP-SWD----LVSAFYSLPLAPLSPT 51
Db 698 GPHDLDGGLLPTPEQTFLPQKRLTP-HPHPHALGPRAWDHGHPLLPASASSSSLLLLAPA 756
Qy 52 SVPISPVSVGRGPDPAHV 70
Db 757 RAPEQPPAPGE-PTPDGRL 774

RESULT 30
US-10-146-728-544
; Sequence 544, Application US/10146728
; Publication No. US20030203437A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C321
; CURRENT APPLICATION NUMBER: US/10/146,728
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-146-728-544

Query Match 20.0%; Score 83.5; DB 12; Length 888;
Best Local Similarity 35.4%; Pred. NO. 8.8;
Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

Qy 1 GTH----SLPRPAAVPLRMQGPAPHVLSFLRP-SWD----LVSAFYSLPLAPLSPT 51
Db 698 GPHDLDGGLLPTPEQTFLPQKRLTP-HPHPHALGPRAWDHGHPLLPASASSSSLLLLAPA 756
Qy 52 SVPISPVSVGRGPDPAHV 70
Db 757 RAPEQPPAPGE-PTPDGRL 774

Search completed: July 4, 2004, 04:29:41
Job time : 18.0221 secs

GenCore version 5.1.1.6
 Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2004, 04:17:04 ; Search time 5.55221 Seconds
 (without alignments)
 1368.668 Million cell updates/sec

Title: US-09-506-079H-11
 Perfect score: 418
 Sequence: 1 GTHSLPRPAAPVFLRMQP.....VGRGPDPAFVAVNLRSYEG 79

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 90 summaries

Database : PIR 78:*
 1: piri:*
 2: pir2:*
 3: pir3:*
 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81.5	19.5	1257	2 S28764	neurocan precursor
2	81	19.4	200	2 H84715	probable phytocyan
3	81	19.4	1006	2 G86292	hypothetical prote
4	80	19.1	1834	1 JDMU1	DNA-directed RNA p
5	80	19.1	1840	2 G85422	hypothetical prote
6	78	18.7	249	2 S72619	hypothetical prote
7	77.5	18.5	217	2 T51031	related to finger
8	77	18.4	503	2 T19319	hypothetical prote
9	76.5	18.3	1110	2 T19673	hypothetical prote
10	76	18.2	356	2 A96826	T8K14.10 imported
11	76	18.2	848	2 S48273	probable transcript
12	75	17.9	189	2 D49600	genome-linked prot
13	75	17.9	1952	2 T48814	hypothetical prote
14	74.5	17.8	894	2 T13029	beta-adaptin homol
15	73	17.5	1520	2 T00273	hypothetical prote
16	72.5	17.3	518	2 F70831	probable pPE prote
17	72	17.2	440	2 A44081	kappa-type opioid
18	72	17.2	1914	2 T42635	tenascin Y precurs
19	71.5	17.1	2441	2 S39161	CREB-binding prote
20	71	17.0	377	2 E87022	probable conserved
21	71	17.0	906	2 A71438	probable resistanc
22	70.5	16.9	291	2 S27721	hypothetical prote
23	70.5	16.9	443	2 T27877	hypothetical prote
24	70.5	16.9	788	1 JVDLHH	DNA-directed DNA p
25	70.5	16.9	1368	2 S53781	neurocan - mouse
26	70	16.7	491	2 S14182	DNA-directed RNA p
27	70	16.7	650	2 S14181	DNA-directed RNA p
28	70	16.7	954	2 E86174	protein F19P19.26
29	70	16.7	2187	2 T30826	nascent polypeptid

RESULT 1

S28764
 neurocan precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text_change 04-Feb-2000
 C:Accession: S28764
 R.Rauch, U.; Karthikeyan, L.; Maurel, P.; Margolis, R.U.; Margolis, R.K.

ALIGNMENTS

30	70	16.7	2318	2 S45306	notch3 protein - h
31	70	16.7	2321	2 S78549	TEA domain-contain
32	69.5	16.6	433	2 T09284	gag polyprotein -
33	69.5	16.6	608	2 A46312	hypothetical prote
34	69.5	16.6	684	2 T25603	caudal-type homeot
35	69	16.5	148	2 T38881	S-layer-like array
36	69	16.5	280	2 A75526	myosin-I binding p
37	69	16.5	1121	2 T02764	ich1 protein - ink
38	69	16.5	1353	2 T00249	proline-rich prote
39	68.5	16.4	401	2 T51407	arabinogalactan-pr
40	68.5	16.4	461	2 T10285	probable spermidin
41	68.5	16.4	554	2 T36545	probable potassum
42	68.5	16.4	1017	2 T31354	CDA peptide synthet
43	68.5	16.4	7463	2 T36248	hybrid proline-ric
44	68	16.3	301	2 J01663	hypothetical prote
45	68	16.3	444	2 R38802	hypothetical prote
46	68	16.3	543	2 S35047	mucin JUL7 - human
47	68	16.3	1206	2 S24407	formin isoform IV
48	68	16.3	1468	2 S11515	formin - mouse
49	68	16.3	1752	1 S26849	DNA-directed RNA p
50	68	16.3	1759	2 T18868	myoblast city prot
51	68	16.3	1859	1 A34092	DNA-directed RNA p
52	68	16.3	1862	2 T29959	DNA-directed RNA p
53	67.5	16.1	123	1 VHWHE	structural protein
54	67.5	16.1	213	1 JMHQ	amelogenin I precu
55	67.5	16.1	222	2 A72683	hypothetical prote
56	67.5	16.1	277	2 T38857	microtubule-associ
57	67.5	16.1	457	2 C95017	peptidase, M20/M25
58	67.5	16.1	457	2 D97890	succinyl-diaminopi
59	67.5	16.1	505	2 A54190	cerebroside-sulfat
60	67.5	16.1	607	2 S74254	homeotic protein s
61	67.5	16.1	827	2 S59121	SOM6 protein - mou
62	67	16.0	190	2 C88560	protein C48B4.12a
63	67	16.0	222	2 H96711	hypothetical prote
64	67	16.0	357	2 S18236	omega secalin prec
65	67	16.0	384	2 H70580	hypothetical prote
66	67	16.0	634	2 T00388	hypothetical prote
67	67	16.0	817	2 S51342	verprolin - yeast
68	66.5	15.9	194	2 A96780	hypothetical prote
69	66.5	15.9	212	2 T35187	merozoite surface
70	66.5	15.9	259	2 S01704	hypothetical prote
71	66.5	15.9	315	2 T06806	proline rich prote
72	66.5	15.9	367	2 T24298	hypothetical prote
73	66.5	15.9	415	2 T00614	hypothetical prote
74	66.5	15.9	464	2 S22697	extensin - Volvox
75	66.5	15.9	522	2 S52216	viral proteinase -
76	66.5	15.9	574	2 F75356	serine/threonine p
77	66.5	15.9	627	2 T26054	hypothetical prote
78	66.5	15.9	1494	2 T14355	protein-tyrosine-p
79	66	15.8	450	2 B70506	hypothetical prote
80	66	15.8	534	2 S21961	proline-rich prote
81	66	15.8	558	2 JC2004	nuclear factor-kap
82	66	15.8	718	2 T29708	hypothetical prote
83	66	15.8	860	2 JC5702	ErB kinase activa
84	66	15.8	868	2 JC5701	ErB kinase activa
85	66	15.8	957	2 T41246	probable ribonucle
86	66	15.8	977	2 S14183	DNA-directed RNA p
87	66	15.8	1447	2 A54100	tumor suppressor p
88	66	15.8	3149	1 Q08E8	BPLF1 protein - hu
89	65.5	15.7	299	2 H98319	hypothetical prote
90	65.5	15.7	357	2 A39384	GDF-1 embryonic gr

J. Biol. Chem. 267, 19536-19547, 1992

A,Title: Cloning and primary structure of neurocan, a developmentally regulated, aggregating proteoglycan core protein from sheep brain

A,Reference number: S28764; PMID:92406907; PMID:1326557

A,Accession: S28764

A,Molecule type: mRNA

A,Residues: 1-1257 <RAW>

A,Cross-references: EMBL:M97161; NID:G205649; PIDN:AAC37679.1; PID:G205650

A,Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; aggrecan; C-type lectin homology; proteoglycan; glycoprotein

A,Keywords: chondroitin sulfate #status predicted <SIG>

F,1-22/Domain: signal sequence #status predicted <SIG>

F,23-1257/Product: neurocan #status predicted <RAW>

F,176-253/Domain: link protein repeat homology <LNK1>

F,274-355/Domain: link protein repeat homology <LNK2>

F,364-366/Region: cell attachment (R-G-D) motif

F,953-984/Domain: EGF homology <EGF>

F,1029-1149/Domain: C-type lectin homology <LCH>

F,1156-1312/Domain: complement factor H repeat homology <FHD>

F,1321-339,737,967,1164/Binding site: carbohydrate (Asn) (covalent) #status predicted

F,372,410/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

F,944/Binding site: chondroitin sulfate (Ser) (covalent) #status experimental

Query Match 19.5%; Score 81.5; DB 2; Length 1257;

Best Local Similarity 35.0%; Pred. No. 3-5;

Matches 28; Conservative 10; Mismatches 25; Indels 17; Gaps 5;

QY 4 SLLPRPAVPLRMQPG---PAHPVLSPLR-----PSWDIVSAFYSILPLAPLS--PT 51

DB 610 SLIPSEALSAVSLQAGSGSDPFPVAMLRAPKMLPLHSTLLVFNWSPILSPASPLPS 669

QY 52 SVP-----ISPVSVGRGPD 67

DB 670 SVPEEQAVRPVSFG-AEDPE 688

RESULT 2

H84715

probable phytoeyanin [imported] - Arabidopsis thaliana

C,Species: Arabidopsis thaliana [mouse-ear cress]

C,Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 21-Oct-2002

C,Accession: H84715

R,Lin, X.: Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;

Reus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, M.

Nature 402, 761-768, 1999

A,Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A,Reference number: A84420; PMID:20083487; PMID:10617197

A,Accession: H84715

A,Status: preliminary

A,Molecule type: DNA

A,Residues: 1-200 <STO>

A,Cross-references: GB:AE002093; NID:G3746072; PIDN:AAC63847.1; GSPDB:GN00139

C,Genetics:

A,Gene: At2g31050

A,Map position: 2

C,Superfamily: stelliacyanin

Query Match 19.4%; Score 81; DB 2; Length 200;

Best Local Similarity 34.3%; Pred. No. 0-47;

Matches 23; Conservative 5; Mismatches 23; Indels 16; Gaps 2;

QY 9 PAAPVPLRMQPGAHPLVLSFLAPSLDLSAFYSLPLAPLSPTSPVPSVSVGRGPDPA 68

DB 132 PVAAPV-----PGVVRPSSFSFSPS-----QSPLAESFVNHPVQYQMGSPAP 175

QY 69 HVAVNL 75

DB 176 HSAASNS 182

RESULT 3

G86292

hypothetical protein F7H2.17 [imported] - Arabidopsis thaliana

C,Species: Arabidopsis thaliana [mouse-ear cress]

A:Gene: rpl1215; RPB1
A:Map position: 4
A:Introns: 28/3; 123/3; 218/3; 271/3; 325/3; 411/1; 440/3; 552/3; 648/2; 734/3; 1755/2;
A:Note: F4B4.70
C:Superfamily: human DNA-directed RNA polymerase II largest chain
C:Keywords: DNA binding; nucleoside transferase; phosphoprotein; tandem repeat; transcrip
F:66-109/Region: zinc finger CCCC motif
F:1531-1812/Region: 7-residue repeats

Query Match 19.1%; Score 80; DB 1; Length 1834;
Best Local Similarity 36.4%; Pred. No. 7.5;
Matches 28; Conservative 3; Mismatches 36; Indels 10; Gaps 3;

QY 7 PRPAAVPVLRMQCP---GPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVSVGCR 62
DB 1730 PSIAVSPSNARLSPASPSPTSPNYSPSPSPSPSYSS---PSSPTSPSPSYSSGA 1785
QY 63 GPD--PDAAHVAVNLSRY 77
DB 1786 SPDYSPSAGYSTPLGTY 1802

RESULT 5
G85422
Hypothetical protein AT4G35800 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C:Accession: G85422
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: G85422
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1840 <STO>
A:Cross-references: GB:NC_001268; NID:G7270532; PIDN:CAB81489.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4G35800
A:Map position: 4
C:Superfamily: human DNA-directed RNA polymerase II largest chain

Query Match 19.1%; Score 80; DB 2; Length 1840;
Best Local Similarity 36.4%; Pred. No. 7.5;
Matches 28; Conservative 3; Mismatches 36; Indels 10; Gaps 3;

QY 7 PRPAAVPVLRMQCP---GPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVSVGCR 62
DB 1736 PSIAVSPSNARLSPASPSPTSPNYSPSPSPSPSYSS---PSSPTSPSPSYSSGA 1791
QY 63 GPD--PDAAHVAVNLSRY 77
DB 1792 SPDYSPSAGYSTPLGTY 1808

RESULT 6
S72619
Hypothetical protein la - anthracnose fungus (Colletotrichum gloeosporioides) retrotrans
C:Species: Glomerella cingulata, Colletotrichum gloeosporioides
C:Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 17-Mar-1999
C:Accession: S72619
R:He, C.; Nourse, J.P.; Kelemu, S.; Irwin, J.A.G.; Manners, J.M.
Mol. Gen. Genet. 252, 320-331, 1996
A:Title: Cg1: a non-LTR retrotransposon with restricted distribution in the fungal phy
A:Reference number: S72619; MUID:96439839; PMID:8842152
A:Accession: S72619
A:Molecule type: DNA
A:Residues: 1-249 <HEA>
A:Cross-references: EMBL:L76172
A:Experimental source: Biotype B, isolate U062
A:Note: in the authors' translation residues 1-10 are not shown
C:Genetics:
A:Mobile element: retrotransposon Cg1

Query Match 18.7%; Score 78; DB 2; Length 249;
Best Local Similarity 34.7%; Pred. No. 1.2;
Matches 25; Conservative 7; Mismatches 26; Indels 14; Gaps 3;

QY 5 LLPR-----PAAVPVPLRMQCPGPAHPVLSFLRPSW---DLVSAFYSLPLAPLSPTSPV 54
DB 24 LVPRWCHPNTNTAQTPTQQRQSPAMASQPAAGSQGPGLLLSSMNLKPTPPPTSLP 83
QY 55 ISPVSVGSGPDP 66
DB 84 PRP-----RGLTP 91

RESULT 7
T51031
related to finger protein XFG 68 [imported] - Neurospora crassa
N:Alternate names: protein B15120.10
C:Species: Neurospora crassa
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T51031
R:Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatur
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T51031
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-217 <SCH>
A:Cross-references: EMBL:AL389900; GSPDB:GN00116; NCSP:B15120.10
A:Experimental source: BAC clone B15120; strain OR74A
C:Genetics:
A:Gene: NCSP:B15120.10
A:Map position: 6

Query Match 18.5%; Score 77.5; DB 2; Length 217;
Best Local Similarity 37.5%; Pred. No. 1.2;
Matches 21; Conservative 5; Mismatches 21; Indels 9; Gaps 2;

QY 7 PRPAAVPVLRMQCPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVSVGCR 62
DB 109 PRPSTASTPRRTCTPTPRP-----PS---TSALVLELVSPSPSSSLIPCSRR 155

RESULT 8
T19319
Hypothetical protein C15H11.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19319
R:Bardill, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19107
A:Accession: T19319
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-503 <WIL>
A:Cross-references: EMBL:Z81035; PIDN:CAB02737.1; GSPDB:GN00023; CESP:C15H11.5
A:Experimental source: clone C15H11
C:Genetics:
A:Gene: CESP:C15H11.5
A:Map position: 5
A:Introns: 53/3; 156/3; 254/3; 292/2; 331/3; 379/3; 437/1

Query Match 18.4%; Score 77; DB 2; Length 503;
Best Local Similarity 32.7%; Pred. No. 3.4;
Matches 17; Conservative 8; Mismatches 17; Indels 10; Gaps 1;

QY 11 AVFVPLRMQCPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVSVGCR 62
DB 298 SIPLMRQEPG-----SGWYLSAWYSLPNVPLVMVMTGVRPAEVAR 339

RESULT 9

T19673
 hypothetical protein C33B4.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T19673
 R:Colles, L.
 submitted to the EMBL Data Library, February 1995
 A:Reference number: Z19160
 A:Accession: T19673
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1110 <ML>
 A:Cross-references: EMBL:Z48367; PIDN:CAA88324.1; GSPDB:GN00020; CESP:C33B4.3
 A:Experimental source: clone C33B4
 C:Genetics:
 A:Gene: CESP:C33B4.3
 A:Map position: 2
 A:Introns: 20/3; 110/3; 144/3; 145/1; 546/1; 653/3; 934/2; 983/3; 1035/1

Query Match 18.3%; Score 76.5; DB 2; Length 1110;
 Best Local Similarity 28.6%; Pred. No. 9.4; Mismatches 12; Indels 15; Gaps 4;
 Matches 24; Conservative 12

QY 3 HSLLPRPAAYVPLRMO-----PGPAHPVLSFLRPSWDLVSFAFYSLP-LAPLSPTSPIS 56
 DB 744 HPSLPRASASTPQIQOQSSIRPPPPPPPHCEPT--MVAVTFPTSTSSVPPPPPLP 801
 QY 57 PVSVGRGPDPP-----AHVAVN 73
 DB 802 PISSGAPPPPPPPPPGGLMVAAS 825

RESULT 10
 A96826
 T8KJ4.10 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: A96826
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Creasy, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:111130712
 A:Accession: A96826
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-356 <STO>
 A:Cross-references: GB:AB005173; NID:G4835761; PIDN:AAD30228.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: T8KJ4.10
 A:Map position: 1

Query Match 18.2%; Score 76; DB 2; Length 356;
 Best Local Similarity 34.7%; Pred. No. 2.9; Mismatches 7; Indels 12; Gaps 3;
 Matches 26; Conservative 7

QY 9 PAAYVPLRMOGPAHPVLSFLRPSWDLVSFAFYSLP-LAPLSPTSPIS 60
 DB 216 PGPVLGPPYSFGSTPTGSIPTSPS-----SGFLPPIVYPPPMAPPSPSVPTSAWCVAK 271
 QY 61 GRGPDPAHVAVNLS 75
 DB 272 PSVDPPIQEAANFA 286

RESULT 11
 S48273

probable transcription factor YBR108w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein YBR0901
 C:Species: Saccharomyces cerevisiae
 C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002
 C:Accession: S48273; S45976; S44688
 R:Mannhaupt, G.; Stucka, R.; Ehle, S.; Vetter, I.; Feldmann, H.
 yeast 10, 1363-1381, 1994
 A:Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.
 A:Reference number: S48255; MUID:95208357; PMID:7900426
 A:Accession: S48273
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-848 <MAN>
 A:Cross-references: EMBL:X78993; NID:G476045; PIDN:CAA55611.1; PID:G476064
 R:Feldmann, H.; Mannhaupt, G.; Schwarze, C.; Vetter, I.
 submitted to the Protein Sequence Database, August 1994
 A:Reference number: S45927
 A:Accession: S45976
 A:Molecule type: DNA
 A:Residues: 1-848 <PE2>
 A:Cross-references: EMBL:Z35977; NID:G536378; PID:G536379; MIPS:YBR108w
 C:Genetics:
 A:Cross-references: SGD:S0000312
 A:Map position: 2R

Query Match 18.2%; Score 76; DB 2; Length 848;
 Best Local Similarity 34.3%; Pred. No. 7.8; Mismatches 28; Indels 10; Gaps 3;
 Matches 23; Conservative 6

QY 13 PVFLRMQGPAPHPV-----LSFLRPSWDLVSA---FYSLP-LAPLSPTSPV---ISPVSVGR 62
 DB 383 PVFVRMQPQPQPMQOQNIPTIEPSLDSTGTFHFVTFDFDAPAPKIDITVDVSS 442
 QY 63 GPDPDAH 69
 DB 443 LPFPPTH 449

RESULT 12
 D49600
 genome-linked protein VPg - soybean dwarf virus
 C:Species: soybean dwarf virus
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Sep-1999
 C:Accession: D49600
 R:Ratthjen, J.P.; Karageorgos, L.E.; Habili, N.; Waterhouse, P.M.; Symons, R.H.
 Virology 198, 671-679, 1994
 A:Title: Soybean dwarf luteovirus contains the third variant genome type in the luteov
 A:Reference number: A49600; MUID:94120742; PMID:8291248
 A:Accession: D49600
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-189 <RAT>
 A:Cross-references: GB:I24049; NID:G436017; PIDN:AAAI7538.1; PID:G436021
 C:Superfamily: potato leaf roll virus genome-linked protein

Query Match 17.9%; Score 75; DB 2; Length 189;
 Best Local Similarity 31.1%; Pred. No. 1.7; Mismatches 22; Indels 8; Gaps 2;
 Matches 19; Conservative 12

QY 2 THSLPRPAAYVPL-----RMQGPAPHPVLSFLRPSWDLVS---AFYSLP-LAPLSPTSV 53
 DB 58 THSCFQRTASVVPREVLSGLYQNASHLMYSRPTWNIRSVSYSSSPRLPFRQV 117
 QY 54 P 54
 DB 118 P 118

RESULT 13
 T48814
 hypothetical protein 1586.220 [imported] - Neurospora crassa
 C:Species: Neurospora crassa
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000


```

Query Match      17.5%  Score 73  DB 2  Length 1520;
Best Local Similarity 35.1%  Pred. No. 30;
Matches 27;  Conservative 3;  Mismatches 25;  Indels 22;  Gaps 4;

Qy  5  LLRPPA-----AYVPLRLMQGPAHVLPSFLRPSNDLVSAFYSLPLAPLPTSPVIS 56
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   710 LLARPpppVQSVGPAVPTFFSMGAALFFAGGGLGMP-----LLPPPLQPPSLPLS 761

```

Query Match	17.5%	Score 73	DB 2	Length 1520
Best Local Similarity	35.1%	Pred: No. 30		
Matches	27	Conservative	3	Mismatches 25; Indels 22; Gaps 4;
Qy 5	LLRPPA-----	AVPVLRMQGP	PAHVLSP	LRFSNDLVSAFYSLPLAPLPTSTVPIS 56
Db 710	LLARPppvqvsv	PAVETPFS	MAALFF	AGGLGMP-----SLPPPLQPPSLPLS 761
Qy 57	--FVSVGRGPD	PDARHVA	71	
Db 762	MGPVL----	EDPETHVA	774	

RESULT 16
F70831
probable PPS protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: F70831
R:Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Ho
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complet
A:Reference number: A70500; MUID:98235987; PMID:9634230
A:Accession: F70831

A:Residues: 1-518 <COL>
A:Cross-references: GB:AL021932; GB:AL123456; NID:g3261527; PIDN:CAAL7410.1; PI
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: PPE

```

Query Match      17.3%  Score 72.5;  DB 2;  Length 518;
Best Local Similarity 30.3%;  Pred. No. 9.8;  Caps 3;
Matches 23;  Conservative 6;  Mismatches 28;  Indels 19;
Qy 1 GTHSLPRPAAVVPVLRMQCPGPAHPVLFSFLRPSNDLVSAFYSLPLA-PLSPPTGVPIPSVPS 59
Db 315 GLTGLVFQPPVVPAP---APDAVVPTV-----LFLAGTATPTTAPASAPA 356
Qy 60 VGRGPDPPDAHVAVNLS 75
Db 357 AGAAGPPGAGTATATS 372

```

RESULT 17
A44081
kappa-type opicoid receptor - human
C:Species: Homo sapiens (man)
C:Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 20-Apr-2000
C:Accession: A44081
E:Xie, G.X.; Miyajima, A.; Goldstein, A.
Proc. Natl. Acad. Sci. U.S.A. 89, 4124-4128, 1992
A:Title: Expression cloning of cDNA encoding a seven-helix receptor from human
A:Reference number: A44081; MUID:9227319; PMID:1315051
A:Accession: A44081
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-440 <GI>
A:Cross-references: GB:M84605; NID:g189391; PID:AAA36395.1; PID:g189392
C:Superfamily: neurokinin 1 receptor
C:Keywords: G protein-coupled receptor; transmembrane protein

```

Query Match      17.2%; Score 72; DB 2; Length 440;
Best Local Similarity 45.2%; Pred. No. 9.1;
Matches 19; Conservative 5; Mismatches 10; Indels 8; Gaps 3;

Qy 2 TSLPRPAAVFVPLRMQGPAPHPVLSTLRPSWDLVSAFYSL 43
   |||||
Db 32 TAGSPAPSWTPSP--RPGPAHP--FLQPPWAV--ALNSL 65

RESULT 18
tenascin Y precursor, variant 206.5K - chicken
C:Species: Gallus gallus (chicken)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T42635
R:Hagios, C.; Koch, M.; Chiquet, M.; Spring, J.; Chiquet-Ehrismann, R.
J. Cell Biol. 134, 1499-1512, 1996
A:Title: Tenascin-Y: a protein of novel domain structure is secreted by differentiated fibroblasts
A:Reference number: 222222; MUID:96427453; PMID:8830777
A:Accession: T42635
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1914 <HAG>
A:Cross-references: EMBL:X99062; NID:gl419545; PIDN:CAA67509.1; PID:gl419546
C:Genetics:
A:Gene: tn-y
C:Keywords: extracellular matrix; heparin binding; heptað repeat
9:1-19/Domain: signal sequence #status predicted <SIG>
9:20-1914/Product: tenascin Y, variant 206.5K #status predicted <MAT>

Query Match      17.2%; Score 72; DB 2; Length 1914;
Best Local Similarity 37.2%; Pred. No. 49;
Matches 32; Conservative 5; Mismatches 25; Indels 24; Gaps 8;

Qy 7 PRPAAVPVPL-----RMQGPAPHPVLSFLRPS--WDLVSAFYSLPLAPLSPTSVPI 55
   |||||
Db 297 PESSAAPASLESFWSPASRSPSPASPV-SPRSPTSPWSPAS-----PQSPILSPAS-PI 349

Qy 56 SPVSVGREGD-POAH-VAVNLSRYEG 79
   |||||
Db 350 SPVL-----PNVPSLHELGVKLSSYNG 371

RESULT 19
S39161
CREB-binding protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 02-Aug-2002
C:Accession: S39161
R:Chirivia, J.C.; Kwok, R.P.S.; Lamb, N.; Hagiwara, M.; Montminy, M.R.; Goodman, R.H.
Nature 365, 855-859, 1993
A:Title: Phosphorylated CREB binds specifically to the nuclear protein CBP.
A:Reference number: S39161; MUID:94019866; PMID:8413673
A:Accession: S39161
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2441 <CHR>
A:Cross-references: GB:S66395; NID:9435854; PIDN:AAB28651.1; PID:9435855
C:Superfamily: transcription coactivator CREB-binding protein; bromodomain homology
F:1112-1169/Domain: bromodomain homology <BRO>

Query Match      17.1%; Score 71.5; DB 2; Length 2441;
Best Local Similarity 36.5%; Pred. No. 73;
Matches 23; Conservative 4; Mismatches 33; Indels 3; Gaps 2;

Qy 4 SLIPRPAAVPVPLRMQGPAPHPVLSTLRPSWDLVSAFYSLPLAPLSPTSVPISPVSGRG 63
   |||||
Db 843 SQVPCFVPTQSPHLPTPPPPASTAAGM--PSLQHTAPGMTFPQFAAPTQ-PSTFVSSGQT 899

Qy 64 PDP 66
   |||
Db 900 PTP 902

```

C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 08-Oct-1999
C;Accession: S33615; S75613; S27721
R;Malakhov, M.P.; Wada, H.; Los, D.A.; Sakamoto, T.; Murata, N.
Plant Mol. Biol. 21, 913-918, 1993
A;Title: Structure of a cyanobacterial gene encoding the 50S ribosomal protein L9.
A;Reference number: S33614; MUID:93222488; PMID:8457083
A;Accession: S33615
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-291 <MA2>
A;Cross-references: GS:DI0716; NID:G217098; PIDN:BAA38818.1; PID:dl002032; PID:G217100
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1992
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shingo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S75613
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-291 <KAN>
A;Cross-references: EMBL:D90912; GB:AB001339; NID:GL653228; PIDN:BA18174.1; PID:dl01890
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 16.9%; Score 70.5; DB 2; Length 291;
Best Local Similarity 28.8%; Pred. No. 8;
Matches 23; Conservative 16; Mismatches 30; Indels 11; Gaps 5;
Qy 4 SLLPRPAVPVPLRMQCPAPVLSFLRPSWD-LVSAFYSLP-LAPLSPT-----SVP 54
Db 145 AIAPEPTLTAPISFPFSP-DPVLSEETPPPAWVNSTFNQFESAPIDSELDQDFATP 203
Qy 55 LSPVSVGRGPD-PDAHVAVN 73
Db 204 ELPLAVEAKPSPSPDMAVS 223

RESULT 23
T27877
hypotheical protein ZK470.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T27877
submitted to the EMBL Data Library, October 1995
A;Description: The sequence of C. elegans cosmid ZK470.
A;Reference number: Z20433
A;Accession: T27877
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-443 <MIN>
A;Cross-references: EMBL:U39651; PIDN:AAA80397.1; CESP:ZK470.5
A;Gene: CESP:ZK470.5
A;Introns: 39/3; 83/3; 109/3; 214/3; 270/3; 332/3; 352/3; 416/2

Query Match 16.9%; Score 70.5; DB 2; Length 443;
Best Local Similarity 31.3%; Pred. No. 13;
Matches 26; Conservative 8; Mismatches 30; Indels 19; Gaps 3;
Qy 7 PRPAVPVPLRMQCPAPVLSFLRPSWDLYSAFYSLPLAPLSPTSVPSVGRGP-- 64
Db 188 PAAAVPAPIMQAP?--PKQASRSSFEVVVALISFDAS-----SSEELSPKKGRL 238

Qy 65 -----DPDAHVAVNLSRYEG 79
Db 239 EIVDHPEDPDWMMARNASGTTG 261

RESULT 24
JDVLHH
DNA-directed DNA polymerase (EC 2.7.7.7) - heron hepatitis virus

C;Species: heron hepatitis virus, HHV
A;Note: host Ardea cinerea (gray heron)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 11-Jun-1999
C;Accession: A30082
R;Spengel, R.; Kalleta, E.P.; Will, H.
J. Virol. 62, 3832-3839, 1988
A;Title: Isolation and characterization of a hepatitis B virus endemic in herons.
A;Reference number: A93037; MUID:88333160; PMID:3418788
A;Accession: A30082
A;Molecule type: RNA
A;Residues: 1-788 <SPR>
A;Cross-references: GB:M22056; NID:G325452; PIDN:AAA45738.1; PID:G325454
C;Superfamily: hepatitis virus DNA-directed DNA polymerase
C;Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 16.9%; Score 70.5; DB 1; Length 788;
Best Local Similarity 40.0%; Pred. No. 25;
Matches 18; Conservative 4; Mismatches 12; Indels 11; Gaps 1;

Qy 31 RPSWDLVSAFYSLPLAPLSPTSVPSIS-----PVSVGRGP 64
Db 444 RISLDLSQAFYHLPLAPSSRLAVSDGKQVYFRKAPMGVGLSP 488

RESULT 25
S52781
neurocan - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 04-Feb-2000
C;Accession: S52781
R;Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.
submitted to the EMBL Data Library, February 1995
A;Description: Amino acid sequence of mouse neurocan and brevican and their different
A;Reference number: S52781
A;Accession: S52781
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1268 <RAU>
A;Cross-references: EMBL:X84727; NID:G758623; PIDN:CAA59216.1; PID:G758630
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology;
F:176-253/Domain: link protein repeat homology <LNK1>
F:274-355/Domain: link protein repeat homology <LNK2>
F:964-995/Domain: EGF homology <EGF>
F:1040-1160/Domain: C-type lectin homology <LCH>
F:1167-1223/Domain: complement factor H repeat homology <PHD>

Query Match 16.9%; Score 70.5; DB 2; Length 1268;
Best Local Similarity 26.1%; Pred. No. 43;
Matches 30; Conservative 11; Mismatches 19; Indels 55; Gaps 7;

Qy 6 LPR-PAAVPVPLRMQCPAH-----PVLSTFLR-----PSWDLVSA 39
Db 604 LRLPSPPPAP---SPGFSEALSNAVLSQASADGSPDFPIVAMLRAPKMLPRSTLVPN 660
Qy 40 FYSLPLAPLS--PTSV-----ISPVSVG-----RGDPDA 68
Db 661 MTPVPLSPASPFLPSWVPEQAVRPVSLGADLETPTTAAAPVEASHRSPDADS 715

RESULT 26
S14182
DNA-directed RNA polymerase (EC 2.7.7.6) largest chain (isoform B2) - soybean (fragment
C;Species: Glycine max (soybean)
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 12-Sep-1997
C;Accession: S14182
R;Dietrich, M.A.; Pronger, J.P.; Guilfoyle, T.J.
Plant Mol. Biol. 15, 207-223, 1990
A;Title: Analysis of the genes encoding the largest subunit of RNA polymerase II in Ara
A;Reference number: S11960; MUID:91355869; PMID:2103447
A;Accession: S14182
A;Molecule type: mRNA
A;Residues: 1-491 <DIE>
A;Cross-references: EMBL:X52493

A;Note: translation of the nucleotide sequence is not complete
C;Genetics:
A;Gene: RP81-B2
C;Intros: 416/2; 440/2; 458/2
C;Superfamily: human DNA-directed RNA polymerase II largest chain
C;Keywords: DNA binding; nucleotidyltransferase; tandem repeat; transcription; zinc finger; 201-474/Region: 7-residue-repeats

Query Match 16.7%; Score 70; DB 2; Length 491;
Best Local Similarity 36.5%; Pred. No. 16;
Matches 23; Conservative 3; Mismatches 29; Indels 8; Gaps 2;

Qy 7 PRPAAVPVLRMQP-----GPAHPVLSFLRPSWDLVSFAFYSILPLAPLSPSTVPISPVSVGR 62
Db 391 PSLAYSPPRLSPSSPYSPPTSNYSPTSFSYSTSPSYSGS---PSSPTYPSPSPSYNSGV 446
Qy 63 SPD 65
Db 447 SPD 449

RESULT 27
SI14181
DNA-directed RNA polymerase (EC 2.7.7.6) largest chain (isoform B1) - soybean (fragment)
C;Species: Glycine max [soybean]
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 12-Sep-1997
C;Accession: SI14181
R;Dietrich, M.A.; Prenger, J.P.; Guilfoyle, T.J.
Plant Mol. Biol. 15, 207-223, 1990
A;Title: Analysis of the genes encoding the largest subunit of RNA polymerase II in Arabidopsis thaliana
A;Reference number: SI1360; MUID: 9135869; PMID: 2103447
A;Accession: SI14191
A;Molecule type: DNA
A;Residues: 1-650 <DIE>
A;Cross-references: EMBL:X52492
A;Note: translation of the nucleotide sequence is not complete
C;Genetics:
A;Gene: RP81-B1
A;Intros: 575/2; 599/2; 617/2
C;Superfamily: human DNA-directed RNA polymerase II largest chain
C;Keywords: DNA binding; nucleotidyltransferase; tandem repeat; transcription; zinc finger; 330-633/Region: 7-residue repeats

Query Match 16.7%; Score 70; DB 2; Length 650;
Best Local Similarity 36.5%; Pred. No. 23;
Matches 23; Conservative 3; Mismatches 29; Indels 8; Gaps 2;

Qy 7 PRPAAVPVLRMQP-----GPAHPVLSFLRPSWDLVSFAFYSILPLAPLSPSTVPISPVSVGR 62
Db 550 PSLAYSPPRLSPSSPYSPPTSNYSPTSFSYSTSPSYSGS---PSSPTYPSPSPSYNSGV 605
Qy 63 SPD 65
Db 606 SPD 608

RESULT 28
E86174
Protein F19p19.26 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana [mouse-ear cress]
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: E86174
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.N.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 815-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: R86141; MUID: 21016719; PMID: 11130712

C:Superfamily: notch protein; ankyrin repeat homology; EGF homology
 F:163-195/Domain: EGF homology <EGF1>
 F:474-505/Domain: EGF homology <EGF>
 F:854-885/Domain: EGF homology <EGF2>
 F:1839-1871/Domain: ankyrin repeat homology <AN1>
 F:1872-1904/Domain: ankyrin repeat homology <AN2>
 F:1906-1938/Domain: ankyrin repeat homology <AN3>
 F:1939-1971/Domain: ankyrin repeat homology <AN4>
 F:1972-2004/Domain: ankyrin repeat homology <AN5>

Query Match 16.7%; Score 70; DB 2; Length 2318;
 Best Local Similarity 37.7%; Pred. No. 97;
 Matches 23; Conservative 2; Mismatches 24; Indels 12; Gaps 1;
 QY 6 LPRNAATVPLEMOCGRHPTLSLRPSMDLVSAFYSLPLAPLSPTSVPISPVSVGRGPD 65
 Db 2162 LLNPVAVPLDWARUPFPAPPGESFL-----LPLAPGQLLNPGAPVSPQERPP 2209
 QY 66 P 66
 Db 2210 P 2210

Search completed: July 4, 2004, 04:21:36
 Job time : 9.55221 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2004, 04:14:33 ; Search time 3.80723 Seconds
(without alignments)
1080.456 Million cell updates/sec

Title: US-09-506-079h-11

Perfect score: 418
Sequence: 1 GTHSLPRRAVPVLRMPQ.....VGRGPDPAHVNLRYEG 79

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 90 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83.5	20.0	888	1 SM6B_HUMAN	Q9h3t3 homo sapien
2	81.5	19.5	1257	1 PCGN_RAT	P50677 rattus norv
3	80	19.1	428	2 FXB2_MOUSE	Q64733 mus musculu
4	80	19.1	1840	1 RPB1_ARATH	P18616 arabidopsis
5	78	18.7	2319	1 NTC3_RAT	Q9r172 rattus norv
6	76	18.2	848	1 YBVB_YEAST	P38266 saccharomyc
7	75	17.9	1046	1 IP2_STRAW	Q32k53 streptomyce
8	74.5	17.8	1255	1 PER2_MOUSE	Q50555 homo sapien
9	73.5	17.6	864	1 YC18_HUMAN	Q5ul22 homo sapien
10	73.5	17.6	887	1 SM6B_RAT	O70141 rattus norv
11	73.5	17.6	1822	1 ZAP3_HUMAN	P49750 homo sapien
12	73	17.5	3503	1 DS_DROME	Q24292 drosophila
13	72	17.2	440	1 NK3R_CAVPO	P30098 cavia porce
14	71.5	17.1	828	1 SOX6_HUMAN	P35712 homo sapien
15	71.5	17.1	2441	1 CBP_MOUSE	P45481 mus musculu
16	70.5	16.9	291	1 YD53_SYNT3	P42350 synochocyst
17	70.5	16.9	485	1 PCDF_HUMAN	Q8n8d1 homo sapien
18	70.5	16.9	788	1 DPOL_HPBHE	P13846 heron hepat
19	70.5	16.9	1175	1 HCN4_RABIT	Q9t665 cryptolagus
20	70.5	16.9	1268	1 PCGN_MOUSE	P55066 mus musculu
21	70.5	16.9	1273	1 SN3A_HUMAN	Q96st3 homo sapien
22	70	16.7	2318	1 NTC3_MOUSE	Q61982 mus musculu
23	70	16.7	2321	1 NTC3_HUMAN	Q9um47 homo sapien
24	69.5	16.6	433	1 TEA3_CHICK	Q90701 gallus gall
25	69.5	16.6	482	1 PCDF_MOUSE	Q9wt01 mus musculu
26	69.5	16.6	802	1 ENAH_MOUSE	Q03173 mus musculu
27	69.5	16.6	889	1 HCN2_HUMAN	Q9ul51 homo sapien
28	69.5	16.6	980	1 RIN3_MOUSE	P59729 mus musculu
29	69	16.5	265	1 CXN1_HUMAN	P47902 homo sapien
30	68.5	16.4	886	1 SM6B_MOUSE	Q54951 mus musculu
31	68.5	16.4	1017	1 KCH4_RAT	Q9r184 rattus norv
32	68	16.3	296	1 CT55_MOUSE	Q9r184 mus musculu
33	68	16.3	703	1 ML51_HUMAN	O15234 homo sapien

RESULT 1

SM6B_HUMAN STANDARD; PRT; 888 AA.

ID SM6B_HUMAN
AC Q9H3T3; Q9NRK9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2003 (Rel. 42, Last annotation update)
DE Semaphorin 6B precursor (Semaphorin Z) (Sema Z).
GN SEMA6B OR SEMA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

ALIGNMENTS

Q05859 mus musculu
Q05860 mus musculu
P16354 schizosacch
P16356 caenorhabdi
Q96j00 homo sapien
Q29325 hepatitis e
Q04632 hepatitis e
Q02817 bos taurus
P50448 mus musculu
P70178 mus musculu
P40645 mus musculu
P80191 oryctolagus
P37370 saccharomyc
P37370 saccharomyc
C00532 homo sapien
Q15027 homo sapien
P09125 elmeria ace
P24152 sorghum bic
Q42201 xenopus lae
Q99594 homo sapien
Q95293 homo sapien
P58152 gallus gall
Q36854 homo sapien
Q35569 rattus norv
P42415 mus musculu
P43146 homo sapien
P70211 mus musculu
P03186 epstein-bar
P20863 mus musculu
Q50K07 drosophila
Q50K07 drosophila
Q50K07 drosophila
Q95583 homo sapien
P30028 duck hepati
P17192 duck hepati
P17193 duck hepati
Q99490 homo sapien
P03162 duck hepati
Q8c253 mus musculu
Q35973 mus musculu
Q40500 saccharomyc
P015643 homo sapien
Q03499 hepatitis e
Q13461 homo sapien
P37697 acetobacter
Q62732 canis fami
P50548 homo sapien
P70459 mus musculu
Q08789 mus musculu
Q9669 drosophila
P28478 turnip yell
P17599 bos taurus
Q60520 mus musculu
Q96qu1 homo sapien
P22165 rhesus papi
P40602 arabidopsis
P35084 dictyosteli
Q8tb24 homo sapien

OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A. (ISOFORM 1).
 RA Kimura T., Ishida H.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21248680; PubMed=11350127;
 RA Correa R.G., Sasahara R.M., Bengtson M.H., Katayama M.L.H.,
 RA Salim A.C.M., Brentani M.M., Sogayar M.C., de Souza S.J.,
 RA Simpson A.J.G.;
 RT "Human semaphorin 6b";
 RL Genomics 73:343-348(2001).
 CC -!- FUNCTION: MAY PLAY A ROLE IN BOTH PERIPHERAL AND CENTRAL NERVOUS
 CC SYSTEM DEVELOPMENT (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Comment=Additional isoforms seem to exist;
 CC Name=1;
 CC IsoId=Q9H3T3-1; Sequence=displayed;
 CC Name=2; Synonyms=6B.1;
 CC IsoId=Q9H3T3-2; Sequence=VSP_006044, VSP_006045;
 CC -!- SIMILARITY: Belongs to the semaphorin family.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB022433; BAB20659.1;
 CC EMBL; AF216389; AAF87861.1;
 CC Genew; HGNC:107399; Sema6b;
 CC InterPro; IPR003659; Flexin-like.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
 KW Developmental protein; Alternative splicing.
 FT SIGNAL 1 25
 FT CHAIN 26 888
 FT DOMAIN 26 603
 FT TRANSMEM 604 624
 FT DOMAIN 625 888
 FT DOMAIN 238 547
 FT DOMAIN 661 674
 FT DOMAIN 750 753
 FT CARBOHYD 74 74
 FT CARBOHYD 155 155
 FT CARBOHYD 167 167
 FT CARBOHYD 291 291
 FT CARBOHYD 386 386
 FT CARBOHYD 441 441
 FT CARBOHYD 462 462
 FT VARSPLIC 471 517
 FT
 FT VARSPLIC 518 888
 FT
 FT CONFLICT 30 30
 FT SEQUENCE 888 AA; 95270 MW; 65FB44D6828C70CB CRC64;
 SQ
 Query Match 20.0%; Score 81.5; DB 1; Length 888;
 Best local similarity 35.4%; Pred. No. 0.88;
 Matches 28; Conservative 7; Mismatches 33; Indels 11; Gaps 5;

QY 1 GTH-----SLLPRAAVPELWQPGAPHLVLSFLRP-SWD-----LVSAFYSLPLAPLSP 51
 Db 698 GPHLDLSGLLTPEQTPLPQKRLPT-HPHPLALGPRADWCHPLPASASSLLLLAPA 756
 QY 52 SVPISPVSVGRGPPDPAHV 70
 Db 757 RAPEQPAPGE-PTPDGRL 774
 RESULT 2
 PGCN FAT
 ID PGCN RAT STANDARD; PRT: 1257 AA.
 AC P55067;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3)
 DE (245 kDa early postnatal core glycoprotein) [Contains: 150 kDa adult
 DE core glycoprotein].
 GN CSPG3 OR NCAN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=92406907; PubMed=1326557;
 RA Rauch U., Karthikeyan L., Maurel P., Margolis R.K.;
 RT "Cloning and primary structure of neurocan, a developmentally
 RT regulated, aggregating chondroitin sulfate proteoglycan of brain.";
 RL J. Biol. Chem. 267:19536-19547(1992).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=94230574; PubMed=7513709;
 RA Friedlander D.R., Milev P., Karthikeyan L., Margolis R.K.,
 RA Margolis R.U., Grumet M.;
 RT "The neuronal chondroitin sulfate proteoglycan neurocan binds to the
 RT neural cell adhesion molecules Ng-CAM/L1/NILE and N-CAM, and inhibits
 RT neuronal adhesion and neurite outgrowth.";
 RL J. Cell Biol. 125:669-680(1994).
 CC -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
 CC development by binding to neural cell adhesion molecules (NG-CAM
 CC and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
 CC acid.
 CC -!- TISSUE SPECIFICITY: Early postnatal and adult brain; not expressed
 CC in kidney, lung, liver and muscle.
 CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND
 CC O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
 CC -!- PTM: TWO ISOFORMS WERE FOUND THAT PROBABLY ARISE BY PROTEOLYTIC
 CC DEGRADATION. THE LARGE ISOFORM IS PREDOMINANT IN EARLY POSTNATAL
 CC BRAIN. THE SMALL ISOFORM IS FOUND IN ADULT BRAIN.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -!- SIMILARITY: Contains 2 EGF-like domains.
 CC -!- SIMILARITY: Contains 2 link domains.
 CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
 CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
 CC -!- SIMILARITY: Belongs to the aggrecan/versican proteoglycan family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M97161; AAC37679.1;
 CC FIR; S28764; S28764.
 CC HSSP; P00740; LEDW.
 CC InterPro; IPR002353; Antifreeze1.
 CC InterPro; IPR000152; Asx_hydroxyl_S.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; AB002345; BAA2804.2; ALT_INIT.
CC Genew; HGNC:8846; PER2.
CC MIM; 603426; -;
CC GO; GO:0007623; P:circadian rhythm; TAS.
CC InterPro; IPR001610; PAC.
CC InterPro; IPR000014; PAS_domain.
CC SMART; SM00086; PAC; 1.
CC SMART; SM00091; PAS; 2.
CC PROSITE; PS01112; PAS; 1.
CC Transcription regulation; Nuclear protein; Repeat; Biological rhythms.
CC KW HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
CC FT DOMAIN 109 146
CC FT DOMAIN 182 248
CC FT DOMAIN 319 389
CC FT DOMAIN 398 438
CC FT DOMAIN 510 513
CC FT DOMAIN 789 806
CC FT DOMAIN 842 979
CC FT PRO-RICH.
CC SQ SEQUENCE 1255 AA; 2AEF2C6BD4B6CB0 CRC64;

Query Match 17.8%; Score 74.5; DB 1; Length 1255;
Best Local Similarity 25.8%; Pred. No. 9.7; Indels 26; Gaps 5;
Matches 30; Conservative 10; Mismatches 10

QY 1 GTHSLPRP-----AAVPLRMQ-----PGPAHPVLSFLRPSMDLVSFAFSLPL 45
DB 863 GTVAAPAPPHASFTVPAPVDLQHQFAVQPPFPAPLAPVAFMLPSYFSGTGNLPQ 922

QY 46 A-PLSP-----TSVP-----ISPVSVGRGDP 66
DB 923 APPPSQPPSHPTLSEMASASQPEPFRSTSPQFCACPATRATPPSANGRASPP 979

RESULT 9
ID YC18 HUMAN STANDARD; PRT; 864 AA.
AC Q9ULK2;
DT 16-OCT-2001 (Rel. 40, Created);
DT 16-OCT-2001 (Rel. 40, Last sequence update);
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein KIAA1218 (Fragment).
GN KIAA1218.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirosewa M., Nomura N.,
RA Chara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 6:337-345 (1999).
CC -!- SIMILARITY: Belongs to the ataxin 7 family.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AB033044; BAA86532.1; -;
KW Hypothetical protein. 1
FT NON_TER 1 226 POLY-THR.
FT DOMAIN 223 226 POLY-SER.
FT DOMAIN 552 664 POLY-SER.
FT DOMAIN 786 790 POLY-SER.
SQ SEQUENCE 864 AA; 92138 MW; D55CB8130E48DA23 CRC64;

Query Match 17.6%; Score 73.5; DB 1; Length 864;
Best Local Similarity 35.5%; Pred. No. 8.2; Indels 35; Gaps 7;
Matches 33; Conservative 4; Mismatches 21

QY 4 SLLPRPAA---VPVPLR-MQGPAPHPVLSFLRPSMDLVSFAFSLPLAPLSP-----T 51
DB 517 SPLSPAPAHITTPVPASVLQP-----FSNP-----SAVY-LPSAIPSSLTSSYIMT 562
QY 52 SVPIPSVSVGRGDPDA-----HVAVNLS 75
DB 563 SAMLSSNAFAVTSPPDSALMSHTTAPPHVAATLS 595

RESULT 10
ID SM6B RAT STANDARD; PRT; 887 AA.
AC O70141;
DT 30-MAY-2000 (Rel. 39, Created);
DT 30-MAY-2000 (Rel. 39, Last sequence update);
DT 16-OCT-2001 (Rel. 40, Last annotation update);
DE Semaphorin 6B precursor (Semaphorin 2) (Sema 2).
GN SEMA6B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=98087397; PubMed=9427525;
RA Kikuchi K., Ishida H., Kimura T.;
RT "Molecular cloning of a novel member of semaphorin family genes,
RT semaphorin Z.";
RL Brain Res. Mol. Brain Res. 51:229-237 (1997).
CC -!- FUNCTION: MAY PLAY A ROLE IN BOTH PERIPHERAL AND CENTRAL NERVOUS
CC -!- SYSTEM DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DEVELOPMENTAL STAGE: DETECTED IN THE FIRST BRANCHIAL ARCH OF
CC EMBRYONIC DAY 11 (Bil) EMBRYO, AND SUBSEQUENTLY IN THE MYOTOMES
CC AND THE DORSAL ROOT GANGLIA IN DEVELOPING SOMITES FROM E11.5
CC THROUGH E13.5, BUT NOT IN THE BRAIN. HOWEVER, AT E15, 18, 21 AND
CC P0, SEMAZ WAS HIGHLY EXPRESSED IN THE BRAIN.
CC -!- SIMILARITY: Belongs to the semaphorin family.
CC -!- SIMILARITY: Contains 1 Sema domain.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AB000776; BAA25687.1; -;
CC InterPro; IPR003659; Plexin-like.
CC InterPro; IPR001627; Sema.
CC Pfam; PF01403; Sema; 1.
CC SMART; SM00423; PSI; 1.
CC SMART; SM00630; Sema; 1.
CC Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
CC Developmental protein.
CC SIGNAL 1 26
CC FT CHAIN 27 887
CC FT DOMAIN 27 605
CC FT TRANSMEM 606 626
CC POTENTIAL.
CC SEMAPHORIN 6B.
CC EXTRACELLULAR (POTENTIAL).
CC POTENTIAL.

```

FT DOMAIN 627 887 CYTOLASMIC (POTENTIAL).
FT DOMAIN 239 549 SEMA.
FT CARBOHYD 75 75 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 887 AA; 95752 MW; 095433F202CD301 CRC64;

Query Match 17.6%; Score 73.5; DB 1; Length 887;
Best Local Similarity 36.0%; Pred. No. 8.4;
Matches 27; Conservative 6; Mismatches 31; Indels 11; Gaps 5;

QY 1 GTH----SLPRPAAVPLRMQGPAPVLSFLRP-SWD----LVSAFYSLPLAPLSPT 51
DB 699 GPHLDGSLLPPTQTLPLQKRLP-TTTPHAHLGPRAWDHSHALLSASTSILLIAHT 757

QY 52 SVPTIS-PVSVGRGPD 65
DB 758 RAPEQPPVPTBSGPE 772

RESULT 11
ZAP3 HUMAN
ID ZAP3 HUMAN STANDARD; PRT; 1822 AA.
AC P49750; P49752; Q9PIV7;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nuclear protein ZAP3 (ZAP113).
GN ZAP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP Rowen L., Madan A., Qin S., Abbasi N., Baradarani L., Birditt B.,
RA Bloom S., Dors M., Dickhoff R., Fleetwood P., Harrison G., James R.,
RA Kaur A., Madan A., Owen M.P., Ratcliffe A., Shaffer T., Hood L.;
RT "Sequencing of human chromosome 14q24.3 region."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 539-847 AND 1397-1822 FROM N.A.
RC TISSUE=Brain;
RX Sherrington R., Rogaeve E.I., Liang Y., Rogaeve E.A., Levesque G.,
RA Ikeda M., Chi H., Lin C., Li G., Holman K., Teuda T., Mar L.,
RA Foncin J.-F., Bruni A.C., Montes M.P., Sorbi S., Rainero I.,
RA Pinessi L., Nee L., Chumakov I., Pollen D., Brookes A.,
RA Sansau P., Polinsky R.J., Wasco W., da Silva H.A.R., Haines J.L.,
RA Pericak-Vance M.A., Tanzi R.E., Roses A.D., Fraser P.E.,
RA Rommens J.M., St George-Hyslop P.H.;
RT "Cloning of a gene bearing missense mutations in early-onset familial
RA Alzheimer's disease."
RL Nature 375:754-760(1995).
CC -! SUBCELLULAR LOCATION: Nuclear (potential).
CC -! CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 1661.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/annouce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AC007956; AAF61275.1; -
CC EMBL; L40403; AAC42008.1; ALT_FRAME.
CC EMBL; L40400; AAC42006.1; -

```

```

KW Nuclear protein.
FT DOMAIN 15 205 PRO-RICH.
FT DOMAIN 382 430 GLN-RICH.
FT DOMAIN 807 1209 ARG-RICH.
FT DOMAIN 1488 1577 ARG-RICH.
FT CONFLICT 621 621 P -> S (IN REF. 2).
FT CONFLICT 1404 1404 T -> I (IN REF. 2).
FT CONFLICT 1821 1821 K -> E (IN REF. 2).
SQ SEQUENCE 1822 AA; 204947 MW; 8B6CB93FE540C7D2 CRC64;

Query Match 17.6%; Score 73.5; DB 1; Length 1822;
Best Local Similarity 35.0%; Pred. No. 18;
Matches 21; Conservative 7; Mismatches 23; Indels 9; Gaps 2;

QY 7 PRPAALPTPVFGSAPPTTYHPPLQSGPSEQVNSK-----APLSKALPYSSFSDDQG 63
DB 471 PRPAALPTPVFGSAPPTTYHPPLQSGPSEQVNSK-----APLSKALPYSSFSDDQG 524

RESULT 12
DS DROME
ID DS DROME STANDARD; PRT; 3503 AA.
AC Q24292; Q9VPS4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dachsous protein precursor (Adherin).
GN DS OR CGI17941.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
RP STAGE.
RC TISSUE=Embryo;
RX Clark H.P., Brentrup D., Schneitz K., Bieber A., Goodman C., Noll M.;
RA MEDLINE=95324813; PubMed=7601355;
RT "Dachsous encodes a member of the cadherin superfamily that controls
RA imaginal disc morphogenesis in Drosophila."
RL Genes Dev. 9:1530-1542(1995).
RN [2]
RP REVISIONS.
RA Noll M.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.P.,
RA George R.A., Lewis S.R., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balwle R.W., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Evansgella C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

```

RA Merkurov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Muskern D.R., Pacht J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sinden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*,"
RL Science 287:2185-2195(2000).
RP REVISIONS.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.B.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.B.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review,"
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC -1- FUNCTION: Involved in morphogenesis. May also be involved in cell
CC adhesion.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: Expressed in embryonic ectoderm. In larvae,
CC expression is restricted to imaginal disks and brain.
CC -1- DEVELOPMENTAL STAGE: Expressed throughout embryogenesis where it
CC is first detected during gastrulation. Also expressed in larvae
CC and adults.
CC -1- SIMILARITY: Contains 27 cadherin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; L08811; AAA9329.2; ...
CC ENBL; AEO03588; AAF51468.3; ALT_INIT.
CC HSSP; P5116; INCU.
CC FlyBase; F3gn000497; ds.
CC GO; GO:0005887; C:integral to plasma membrane; ISS.
CC GO; GO:0008014; F:calcium-dependent cell adhesion molecule ac. . . ; ISS.
CC GO; GO:0016339; P:calcium-dependent cell-cell adhesion; ISS.
CC GO; GO:0008283; P:cell proliferation; IMP.
CC GO; GO:0009304; P:cellular morphogenesis during differentiation; IMP.
CC GO; GO:0045317; P:equator specification; IMP.
CC GO; GO:0045198; P:establishment of epithelial cell polarity; IMP.
CC GO; GO:0018149; P:protein-protein cross-linking; IPI.
CC InterPro; IPR002126; Cadherin.
CC InterPro; IPR000233; Cadherin_C-term.
CC Pfam; PF00028; cadherin; 26.
CC Pfam; PF01049; Cadherin_C-term; 1.
CC PRINTS; PR00205; CADHERIN.
CC SMART; SM00112; CA; 25.
CC PROSITE; PS00232; CADHERIN_1; 20.
CC PROSITE; PS00269; CADHERIN_2; 27.
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium; Calcium-binding;
KW Repeat; Signal; Developmental protein.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 3503 DACHSOUTS PROTEIN.
FT DOMAIN 21 3045 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 3046 3066 POTENTIAL.
FT DOMAIN 3067 3503 CYTOPLASMIC (POTENTIAL).
FT

22 121 CADHERIN 1.
23 122 CADHERIN 2.
24 123 CADHERIN 3.
25 124 CADHERIN 4.
26 125 CADHERIN 5.
27 126 CADHERIN 6.
28 127 CADHERIN 7.
29 128 CADHERIN 8.
30 129 CADHERIN 9.
31 130 CADHERIN 10.
32 131 CADHERIN 11.
33 132 CADHERIN 12.
34 133 CADHERIN 13.
35 134 CADHERIN 14.
36 135 CADHERIN 15.
37 136 CADHERIN 16.
38 137 CADHERIN 17.
39 138 CADHERIN 18.
40 139 CADHERIN 19.
41 140 CADHERIN 20.
42 141 CADHERIN 21.
43 142 CADHERIN 22.
44 143 CADHERIN 23.
45 144 CADHERIN 24.
46 145 CADHERIN 25.
47 146 CADHERIN 26.
48 147 CADHERIN 27.
49 148 V -> I (IN REF. 1).
50 149 R -> S (IN REF. 1).
51 150 G -> S (IN REF. 1).
52 151 G -> S (IN REF. 1).
53 152 V -> I (IN REF. 1).
54 153 V -> I (IN REF. 1).
55 154 P -> A (IN REF. 1).
56 155 A -> S (IN REF. 1).
57 156 S -> T (IN REF. 1).
58 157 R -> Q (IN REF. 1).
59 158 R -> G (IN REF. 1).
60 159 SQ SEQUENCE 3503 AA; 379774 MW; 975B09F059F7EEF5 CRC64;
61
62 Query Match 17.5%; Score 73; DB 1; Length 3503;
63 Best Local Similarity 32.1%; Pred. No. 40;
64 Matches 26; Conservative 13; Mismatches 26; Indels 16; Gaps 5;
65
66 QY 9 PAAVDPVLMOP----GPAH----PVLSFLRPSWDLVSAF---YSLPLAPLSPTSVPTSP 57
67 DB 3411 PRANLPKRLPPLSLAPLHLPSPICGHEAGSGFTSSAMSPSPSLPLATKSPSPSP 3470
68 QY 58 VSVRGPPDPDAHVA-VNLSRY 77
69 DB 3471 LGAG-----PPTHLPVSLPRH 3487
70
71 RESULT 13
72 NK3R_CAVPO
73 ID NK3R_CAVPO STANDARD; PRT; 440 AA.
74 AC P30098; Q925R4;
75 DT 01-APR-1993 (Rel. 25, Created)
76 DT 01-APR-1993 (Rel. 25, Last sequence update)
77 DT 15-MAR-2004 (Rel. 43, Last annotation update)
78 DE Neuromedin K receptor (NK3R) (Neurokinin B receptor) (NK-3 receptor)
79 DE (NK-3R) (tachykinin receptor 3).
80 GN TACR3.
81 OS Cavia porcellus (Guinea pig).
82 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
83 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
84 OX NCBI_TaxID=10141;
85 RN [1]
86 RP SEQUENCE FROM N.A.
87 RC TISSUE=Placenta;
88 RX MEDLINE=92237319; PubMed=1315051;
89 Xie G.-X., Miyajima A., Goldstein A.;
90 "Expression cloning of cDNA encoding a seven-helix receptor from

human placenta with affinity for opioid ligands.";
Proc. Natl. Acad. Sci. U.S.A. 89:4124-4128(1992).
[2]
RP SHOWS THAT IT IS NOT AN OPIOID RECEPTOR.
RX MEDLINE=97103087; PubMed=8947459;
RA Donaldson L.P., Haskell C.A., Hanley M.R.;
RT "Functional characterization by heterologous expression of a novel
RL cloned tachykinin peptide receptor.";
RL Biochem. J. 320:1-5(1996).
[3]
RN SEQUENCE FROM N.A., AND SHOWS THAT IT IS NOT A HUMAN SEQUENCE.
RA Page N.M., Bell N.J.;
RP "The human tachykinin NK1 (short form) and tachykinin NK4 receptor: a
RT reappraisal.";
RL Eur. J. Pharmacol. 437:27-30(2002).
[4]
RN SEQUENCE FROM N.A.
RA Stumm R.K., Derst C., Schaefer M.K.H., Weihe E.;
RT "Cloning and sequencing of guinea pig NK3 receptor.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: This is a receptor for the tachykinin neuropeptide
CC neuromedin K (neurokinin B). It is associated with G proteins
CC that activate a phosphatidylinositol-calcium second messenger
CC system.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: The anchoring of this receptor to the plasma membrane is
CC probably mediated by the palmitoylation of a cysteine residue.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Highest to other tachykinins receptors.
CC -!- CAUTION: Was originally (Ref.1) thought to be a kappa-type opioid
CC receptor and to originate from human. Ref.2 showed that it is a
CC tachykinin receptor and was termed NK-4R. Ref.3 shows that it is
CC from guinea pig and is NK-3R.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M84605; AAA36395.1; -;
DR EMBL; AF426173; AAL78507.1; -;
DR EMBL; AF255393; AAK49193.1; -;
DR PIR; A44081; A44081.
DR HSSP; P02699; 1P98.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0004985; F:tachykinin receptor activity; TAS.
DR GO; GO:0004888; F:transmembrane receptor activity; TAS.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
DR GO; GO:0007609; P:mechanosensory perception; TAS.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOODPSN.
DR PROSITE; PS00237; G-PROTEIN RECP_F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Lipoprotein;
KW Palmitate.
FT DOMAIN 1 59 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 60 82 1 (POTENTIAL).
FT DOMAIN 83 92 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 93 114 2 (POTENTIAL).
FT DOMAIN 115 134 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 135 156 3 (POTENTIAL).
FT DOMAIN 157 176 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 177 197 4 (POTENTIAL).
FT DOMAIN 198 220 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 221 245 5 (POTENTIAL).
FT DOMAIN 246 274 6 (POTENTIAL).
FT TRANSMEM 275 296 7 (POTENTIAL).
FT DOMAIN 297 309 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 310 334 7 (POTENTIAL).

FT DOMAIN 335 440 CYTOPLASMIC (POTENTIAL).
FT DISULFID 133 208 BY SIMILARITY.
FT CARBOHYD 24 7 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 349 349 S-palmitoyl cysteine (By similarity).
FT CONFLICT 59 59 A -> R (IN REF. 4).
SQ SEQUENCE 440 AA; 544BB3A71CC6C143 CRC64;

Query Match 17.2%; Score 72; DB 1; Length 440;
Best Local Similarity 45.2%; Pred. NO. 5.6; Mismatches 8; Gaps 3;
Matches 19; Conservative 5;

QY 2 THSLLRPAAPVPLRMQGPAPVLSFLRPSWDLVSAYSL 43
DB 32 TASPSPAPSWTPSP---RGPAPP---FLQPPWAV--ALMSL 65

RESULT 14
SOX6 HUMAN
ID _SOX6 HUMAN STANDARD; PRT; 828 AA.
AC P35712; Q9BXQ3; Q9BXQ4; Q9BXQ5; Q9H018;
DT 01-JUN-1994 (Rel. 29, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcription factor SOX-6.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RC TISSUE=Lymphocytes, and Myoblasts;
RX MEDLINE=21153434; PubMed=11255018;
RA Cohen-Barak O., Hagiwara N., Arit M.F., Horton J.P., Brilliant M.H.;
RT "Cloning, characterization and chromosome mapping of the human SOX6
RL Gene 265:157-164(2001).
[2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Testis;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansoerge W., Boecker M., Bloecker H., Bauersachs S., Slum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA Mewes H.-W., Ottenwaeider B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
PL Genome Res. 11:422-435(2001).
[3]
RP SEQUENCE OF 632-685 FROM N.A.
RX MEDLINE=92310993; PubMed=1614875;
RA Denny P., Swift S., Brand N., Dabhadre N., Barton P., Ashworth A.;
RT "A conserved family of genes related to the testis determining gene,
RT SRY.";
RL Nucleic Acids Res. 20:2887-2887(1992).
CC -!- FUNCTION: Binds specifically to the DNA sequence 5'-AACAAAT-3'.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC SVent-Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=P35712-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P35712-2; Sequence=VSP_002196;
CC Name=3;
CC IsoId=P35712-3; Sequence=VSP_002197;
CC -!- TISSUE SPECIFICITY: Expressed in a wide variety of tissues, most
CC abundantly in skeletal muscle.
CC -!- SIMILARITY: Contains 1 HMG box domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: AF309034; AAK26115.1; -
 DR EMBL: AF309476; AAK26243.1; -
 DR EMBL: AF309471; AAK26243.1; JOINED.
 DR EMBL: AF309472; AAK26243.1; JOINED.
 DR EMBL: AF309473; AAK26243.1; JOINED.
 DR EMBL: AF309474; AAK26243.1; JOINED.
 DR EMBL: AF309475; AAK26243.1; JOINED.
 DR EMBL: AF309476; AAK26244.1; -
 DR EMBL: AF309471; AAK26244.1; JOINED.
 DR EMBL: AF309472; AAK26244.1; JOINED.
 DR EMBL: AF309473; AAK26244.1; JOINED.
 DR EMBL: AF309474; AAK26244.1; JOINED.
 DR EMBL: AF309475; AAK26244.1; JOINED.
 DR EMBL: AL136780; CAB66714.1; ALT_INIT.
 DR EMBL: X65663; CAA46614.1; -
 DR HSP: Q05066; LHRV.
 DR Genew: HGNC:16421; SOX6.
 DR MIM: 607257; -
 DR GO: GO:0005634; C:nucleus; NAS.

DR GO: GO:0003700; P:transcription factor activity; NAS.
 DR GO: GO:0006325; P:establishment and/or maintenance of chromatin; NAS.
 DR GO: GO:0007507; P:heart development; ISS.
 DR GO: GO:0007517; P:muscle development; NAS.
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
 DR InterPro: IPR000910; HMG_12_box.
 DR Pfam: PF00505; HMG_box; 1.
 DR SMART: SM00398; HMG; 1.
 DR PROSITE: PS00118; HMG_BOX_2; 1.
 DR DNA-binding; Nuclear protein; Transcription regulation; Coiled coil;
 KW Alternative splicing.

FT DOMAIN 184 262
 FT DNA_BIND 621 689 HMG_BOX.
 FT DOMAIN 240 243 POLY-GLN.
 FT DOMAIN 280 285 POLY-ALA.
 FT DOMAIN 313 317 POLY-ALA.
 FT DOMAIN 514 518 POLY-GLN.
 FT VARSPPLIC 327 367 Missing (in isoform 2).
 FT VARSPPLIC 327 367 /FTId=VSP_002196.
 FT VARSPPLIC 579 598 Missing (in isoform 3).
 FT VARSPPLIC 579 598 /FTId=VSP_002197.
 FT CONFLICT 477 477 S -> SLGKRSQHQEYIE (IN REF. 2).
 FT CONFLICT 633 633 K -> R (IN REF. 3).
 SQ SEQUENCE 828 AA; 91893 MW; 58CA7CDEA811D5D CRC64;

Query Match 17.1%; Score 71.5; DB 1; Length 828;
 Best Local Similarity 34.8%; Pred. No. 12;
 Matches 23; Conservative 13; Mismatches 27; Indels 3; Gaps 2;

Qy 10 AAVPVLPMQGPAPVLSFLRPSWDLVSFYSLSPLAPUSPTVSPV--SVGRGPD 67

Db 421 AAQPLNLSRPKTAEPVKSPSTQNLFPASKTSPVNLNPKSIP-SPIGGSLRGSSLD 479

Qy 68 ARVAVN 73

Db 480 ILGSLN 485

RESULT 15

CBP_MOUSE

ID CBP_MOUSE

AC P45481; STANDARD; PRT; 2441 AA.

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE CREB-binding protein (BC 2.3.1.48).

GN CREBBP OR CBP.

OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=94019866; PubMed=8413673;
 RA Chivria J.C., Kwok R.P.S., Lamb N., Hagiwara M., Montminy M.R.,
 RA Goodman R.H.;
 RT "Phosphorylated CREB binds specifically to the nuclear protein CBP.";
 RL Nature 365:855-859(1993).
 RP [2]
 RP INTERACTION WITH NCOA3.
 RX MEDLINE=97336097; PubMed=9192892;
 RA Torchia J., Rose D.W., Inostroza J., Kamei Y., Westin S., Glass C.K.,
 RA Rosenfeld M.G.;
 RT "The transcriptional co-activator p/CIP binds CBP and mediates
 RT nuclear-receptor function.";
 RL Nature 387:677-684(1997).
 CC !- FUNCTION: Acetyltransferase enzyme. Acetylates histones, giving a
 CC specific tag for transcriptional activation. Also acetylates non-
 CC histone proteins, like NCOA3 coactivator. Mediates CAMP-gene
 CC regulation by binding specifically to phosphorylated CREB protein.
 CC CBP as coactivator, augments the activity of phosphorylated CREB
 CC to activate transcription of CAMP-responsive genes (By
 CC similarity).
 CC !- CATALYTIC ACTIVITY: Acetyl-CoA + histone = CoA + acetyl-histone.
 CC !- SUBUNIT: Interacts with SMAD1, SMAD2, SMAD3, PCAF and PML. The
 CC TAZ-type 1 domain interacts with HIF1A. Found in a complex
 CC containing NCOA2, NCOA3, IKK2, IKK3 and IKK3G. Interacts with
 CC NCOA6 coactivator. Probably part of a complex with Hif1a and
 CC EP300 (By similarity).
 CC !- SUBCELLULAR LOCATION: Nuclear.
 CC !- SIMILARITY: Contains 1 bromodomain.
 CC !- SIMILARITY: Contains 1 ZZ-type zinc finger.
 CC !- SIMILARITY: Contains 2 TAZ-type zinc fingers.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: S66385; AAB28651.1; -
 DR PIR: S39161; S39161.
 DR PDB: 1F81; 18-OCT-00.
 DR PDB: 1JUS; 03-OCT-01.
 DR PDB: 1KEH; 06-FEB-02.
 DR PDB: 1KDX; 25-NOV-98.
 DR PDB: 1L8C; 24-APR-02.
 DR TRANSFAC: T01318; -
 DR MGD: MGI:1098280; Crebbp.
 DR GO: GO:0000123; C:nucleus acetyltransferase complex; IDA.
 DR GO: GO:0005634; C:nucleus; IDA.
 DR GO: GO:0004403; F:histone acetyltransferase activity; IDA.
 DR GO: GO:0005515; F:protein binding; IPI.
 DR GO: GO:0003713; F:transcription co-activator activity; IDA.
 DR GO: GO:0045449; P:regulation of transcription; IDA.
 DR InterPro: IPR001487; Bromodomain.
 DR InterPro: IPR003101; KIX.
 DR InterPro: IPR000197; TAZ finger.
 DR InterPro: IPR000433; znf_ZZ.
 DR Pfam: PF00439; bromodomain; 1.
 DR Pfam: PF02172; KIX; 1.
 DR Pfam: PF02135; zf-TAZ; 2.
 DR Pfam: PF00569; ZZ; 1.
 DR PRINTS: PR00503; BROMODOMAIN.
 DR SMART: SM00297; BROMO; 1.
 DR SMART: SM00551; znf_TAZ; 2.
 DR SMART: SM00291; znf_ZZ; 1.
 DR PROSITE: PS00633; BROMODOMAIN_1; 1.

DR PROSITE; P850014; BROMODOMAIN 2; 1.
DR PROSITE; P850134; ZF_TAZ; 2.
DR PROSITE; P850135; ZF_ZZ_1; 1.
DR PROSITE; P850135; ZF_ZZ_2; 1.
KW Bromodomain; Zinc-finger; Repeat; 3D-structure.
FT ZN_FING 346 432 TAZ-TYPE 1.
FT DOMAIN 1104 1176 BROMODOMAIN.
FT ZN_FING 1702 1745 ZF-TYPE.
FT ZN_FING 1766 1847 TAZ-TYPE 2.
FT DOMAIN 1062 1065 POLY-GLU.
FT DOMAIN 1556 1563 POLY-GLU.
FT DOMAIN 1944 1949 POLY-PRO.
FT DOMAIN 1968 1971 POLY-GLN.
FT DOMAIN 2082 2086 POLY-GLN.
FT DOMAIN 2200 2216 POLY-GLN.
FT DOMAIN 2296 2299 POLY-GLN.
SQ SEQUENCE 2441 AA; 265474 MW; 0ABBO28C3112F419 CRC64;

Query Match 17.1%; Score 71.5; DB 1; Length 2441;
Best Local Similarity 36.5%; Pred. No. 39;
Matches 23; Conservative 4; Mismatches 33; Indels 3; Gaps 2;

QY 4 SLLPRAAVVPLRMQPGAHVPLSLRPSWDLVSAFYSLPLAPLSPTSPVPSVGVGRG 63
Db 843 SLLPRAAVVPLRMQPGAHVPLSLRPSWDLVSAFYSLPLAPLSPTSPVPSVGVGRG 63

QY 64 PDP 66
Db 900 PTP 902

RESULT 16
YD53 SYNY3
ID YD53 SYNY3 STANDARD; PRT; 291 AA.
AC P42350;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein slr1353.
GN SLR1353.
OS Synechocystis sp. (strain PCC 6803)
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OC NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9222488; PubMed=8467083;
RA Malakhov M.P., Wada H., Los D.A., Sakamoto T., Murata N.;
RT "Structure of a cyanobacterial gene encoding the 50S ribosomal
protein L9."
RL Plant Mol. Biol. 21:913-918(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugiyama M., Sasaki M., Kimura T.,
RA Hosouchi I., Matsuno A., Muraki A., Nakazaki N., Nario K., Okumura S.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
Synechocystis sp. strain PCC6803. II. Sequence determination of the
entire genome and assignment of potential protein-coding regions."
RL DNA Res. 3:109-136(1996).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D10716; BAA38818.1; -

DR EMBL; D90912; BAA18174.1; -.
DR FIR; S33615; S27721.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 291 AA; 31283 MW; 5AB7E3DD03C36390 CRC64;

Query Match 16.9%; Score 70.5; DB 1; Length 291;
Best Local Similarity 28.8%; Pred. No. 5.1;
Matches 23; Conservative 16; Mismatches 30; Indels 11; Gaps 5;

QY 4 SLLPRAAVVPLRMQPGAHVPLSLRPSWDLVSAFYSLPLAPLSPTSPVPSVGVGRG 54
Db 145 AIAPEFTLTTPAPISPPSP-DFVLSLEPTPPAMVNTFNQPEESAPIDSLELDQDFATP 203

QY 55 ISPVSVGRGPD-PDAHVAVN 73
Db 204 ELPLAVEAKPDSPEPDMAVS 223

RESULT 17
PCD7 HUMAN
ID PCD7 HUMAN STANDARD; PRT; 485 AA.
AC Q8N8D1; Q8GAK8; Q9Y6D7;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Programmed cell death protein 7 (ES18) (HES18).
GN PDCD7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kanihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEO human cDNA sequencing project."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 269-485 FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=99157097; PubMed=10037816;
RA Park E.J., Kim J.H., Seong R.H., Kim C.G., Park S.D., Hong S.H.;
RT "Characterization of a novel mouse cDNA, ES18, involved in apoptotic
cell death of T-cells."
RL Nucleic Acids Res. 27:1524-1530(1999).
RN [3]
RP SEQUENCE OF 340-485 FROM N.A.
RC TISSUE=Ductenium;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.M., Murthy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Promotes apoptosis when overexpressed (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
DR EMBL; AK096970; BAC04915.1; -
DR EMBL; AF083930; RAD20241.1; -
DR EMBL; BC016992; ARI16992.2; -
DR EMBL; BT007395; AAP36059.1; -
DR GenBank; HGNC:8767; PDCD7.
DR MIM; 608138; -
KW Apoptosis; Nuclear protein; Coiled coil; Repeat.
FT DOMAIN 4 129
FT PRO-RICH.
FT DOMAIN 236 278
FT ARG/GLU-RICH.
FT DOMAIN 232 335
FT COILED COIL (POTENTIAL).
FT DOMAIN 362 411
FT COILED COIL (POTENTIAL).
FT CONFLICT 292 292 Q -> GE (IN REF. 2).
FT CONFLICT 303 303 S -> A (IN REF. 2).
FT SEQUENCE 485 AA; 54699 MW; FF907D6D2187832A CRC64;
SQ
Query Match 16.9%; Score 70.5; DB 1; Length 485;
Best Local Similarity 33.8%; Pred. No. 8.8;
Matches 27; Conservative 3; Mismatches 29; Indels 21; Gaps 5;
QY 6 LRPAAVPLRMQGPDPHVLVLS--PLRSDVL-----VSAPVSLPLAPLS 49
DB 32 LPSP-AFPPLPQRPQGF-PFGASAPFLQPLALQPRASASRGGGAGAFYVPPPLP 89
QY 50 PTVSV---PISPVSVSGRPP 66
DB 90 PPPQCRPPFGTAGERPP 109
RESULT 18
DPOL HPBHE
ID DPOL HPBHE STANDARD; PRT; 788 AA.
AC P13826;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN P.
OS Heron hepatitis B virus.
OC Viruses; Retroviral viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=28300;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88333160; PubMed=3418789;
RA Sprengel R., Kaleta E.F., Will H.;
RT Isolation and characterization of a hepatitis B virus endemic in
RT herons.
RL J. Virol. 62:3832-3839(1988).
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphonooxester.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC

CC or send an email to license@isb-sib.ch.
CC
DR EMBL; M22056; AAA45738.1; -
DR PIR; A30082; JDLVHH
DR InterPro; IPR001462; DNaPol_viral_C.
DR InterPro; IPR000201; DNaPol_viral_N.
DR InterPro; IPR000477; RVTS.
DR Pfam; PF00336; DNaPol_viral_C; 1.
DR Pfam; PF00242; DNaPol_viral_N; 1.
DR Pfam; PF00078; RVT; 1.
DR ProDom; PD000814; DNaPol_viral_C; 1.
KW Transferase; RNA-directed DNA polymerase; DNA-directed DNA polymerase;
KW Hydrolase; Nuclease; Endonuclease; DNA replication; DNA-binding.
SQ SEQUENCE 788 AA; 90070 MW; FB44F3875EADP44 CRC64;
Query Match 16.9%; Score 70.5; DB 1; Length 788;
Best Local Similarity 40.0%; Pred. No. 15;
Matches 18; Conservative 4; Mismatches 12; Indels 11; Gaps 1;
QY 31 RPSWDLVSAPVSLPLAPLSPTSPIS-----PVSVGRGP 64
DB 444 RISLDSQAFYHLPLAPASSRLSRLAVSDGKQVYFRKAPMGVGLSP 488
RESULT 19
HCN4 RABIT
ID HCN4 RABIT STANDARD; PRT; 1175 AA.
AC QSTV667; QSTU35;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Potassium/sodium hyperpolarization-activated cyclic nucleotide-gated
DE channel 4 (Hyperpolarization-activated cation channel 4) (HAC-4).
DE HCN4 OR HAC4.
GN HCN4 RABIT
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE=Heart atrium;
RX MEDLINE=92230313; PubMed=10212270;
RA Ishii T.M., Takano M., Xie L.-H., Noma A., Ohmori H.;
RT "Molecular characterization of the hyperpolarization-activated cation
RT channel in rabbit heart sinoatrial node."
RL J. Biol. Chem. 274:12835-12839(1999).
RN [2]
RP SEQUENCE OF 261-381 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=9345217; PubMed=10400919;
RA Shi W., Wymore R., Yu H., Wu J., Wymore R.T., Pan Z., Robinson R.B.,
RA Dixon J.E., McKinnon D., Cohen I.S.;
RT "Distribution and prevalence of hyperpolarization-activated cation
RT channel (HCN) mRNA expression in cardiac tissues."
RL Circ. Res. 85:1-6(1999).
CC -!- FUNCTION: Hyperpolarization-activated ion channel with very slow
CC activation and inactivation exhibiting weak selectivity for
CC potassium over sodium ions. May contribute to the native pacemaker
CC currents in heart (if) and in neurons (lh). Activated by CAMP. May
CC mediate responses to sour stimuli.
CC -!- SUBUNIT: The potassium channel is probably composed of a homo- or
CC heterotetrameric complex of pore-forming subunits.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Highly expressed in the heart sinoatrial node
CC (SAN). Not detected in atrium, ventricle, forebrain or cerebellum.
CC Detected at very low levels in total brain.
CC -!- DOMAIN: The segment S4 is probably the voltage-sensor and is
CC characterized by a series of positively charged amino acids at
CC every third position.
CC -!- MISCELLANEOUS: Inhibited by extracellular cesium ions.
CC -!- SIMILARITY: Belongs to the potassium channel family. HCN
CC subfamily.
CC -!- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.

FT DISULFID 279 354 BY SIMILARITY.
 FT DISULFID 303 324 BY SIMILARITY.
 FT DISULFID 964 975 BY SIMILARITY.
 FT DISULFID 969 984 BY SIMILARITY.
 FT DISULFID 986 995 BY SIMILARITY.
 FT DISULFID 1040 1051 BY SIMILARITY.
 FT DISULFID 1068 1160 BY SIMILARITY.
 FT DISULFID 1136 1152 BY SIMILARITY.
 FT DISULFID 1167 1210 BY SIMILARITY.
 FT DISULFID 1196 1223 BY SIMILARITY.
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 742 742 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 978 978 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1175 1175 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1268 AA; 137200 MW; 3014E8E202A2FAEC CRC64;
 Query Match 16.9%; Score 70.5; DB 1; Length 1268;
 Best Local Similarity 26.1%; Pred. No. 24;
 Matches 30; Conservative 11; Mismatches 19; Indels 55; Gaps 7;
 QY 6 LPR-PAAPVPLRMQGPAAH-----PVLSFLR-----PSMDLVA 39
 DB 604 LPRLPSEAP--SFGPSEALSAVLSQASSADGSPDFPIVAMLRAPKWLPRSTLVN 660
 QY 40 FVSLDLAPLS--PTSV--ISPVSVG-----RGSDPDA 68
 DB 661 MTPVPLSPASPLPSWVPEQAVPSLGAEDLETPTQTIAAPVASHRSPDADS 715
 RESULT 21
 ID SN3A HUMAN STANDARD; PRT; 1273 AA.
 AC Q96ST3; Q8N8N4; Q8NC83; Q8WV18;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Paired amphipathic helix protein Sin3a.
 GN SIN3A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Guo J.H., Yu L.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-1026 FROM N.A.
 RC TISSUE=Brain, and Teratocarcinoma;
 RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
 RA Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatsuma M., Murakawa K., Kanehori K., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
 RT "NEO human cDNA sequencing project."
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 857-1273 FROM N.A.
 RC TISSUE=Uterus;
 RX MEDLINE=22388457; PubMed=12477932;
 RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M.J., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.; Smailus D.E.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: Acts as a transcriptional repressor. Interacts with MX11
 CC to repress MYC responsive genes and antagonize MYC oncogenic
 CC activities. Also interacts with MAB-MAX heterodimers by binding to
 CC MAD. The heterodimer then represses transcription by tethering
 CC SIN3A to DNA (By similarity).
 CC -!- SUBUNIT: Interacts with MX11 and MAD (By similarity).
 CC -!- SIMILARITY: Contains 3 PAH (paired amphipathic helix) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF18569; AAP97288.1; -
 CC EMBL; AK027559; BAB55197.1; -
 CC EMBL; AK074903; BAC11280.1; -
 CC EMBL; AK096477; BAC04801.1; -
 CC EMBL; BC018973; AAH18973.1; -
 CC EMBL; BC018973; AAH18973.1; -
 CC Genew; HGNC:19353; SIN3A.
 CC MIN; 607776; -
 CC InterPro; IPR003822; PAH.
 CC Pfam; PF02671; PAH; 3.
 KW Transcription regulation; Repressor; Repeat; Nuclear protein.
 FT REPEAT 141 187 PAH 1.
 FT REPEAT 322 381 PAH 2.
 FT REPEAT 477 523 PAH 3.
 FT CONFLICT 216 216 I -> T (IN REF. 2; BAC04801).
 FT CONFLICT 1009 1009 Q -> R (IN REF. 2; BAC11280).
 FT CONFLICT 1247 1248 TT -> NN (IN REF. 1).
 SQ SEQUENCE 1273 AA; 145174 MW; E6A329BE0ED84CD CRC64;
 Query Match 16.9%; Score 70.5; DB 1; Length 1273;
 Best Local Similarity 26.2%; Pred. No. 24;
 Matches 22; Conservative 8; Mismatches 37; Indels 17; Gaps 3;
 QY 2 THSLLPRAVP-----VPLEMQGPAPVLSFLRPSWDLVSAFYSLPLAPLSPT 51
 DB 213 THGQPQPPPPQHPSPQSAQAPAPAPQPPPAKVSKEPS--QLQAH-----TPASQ 265
 QY 52 SVPIPSVSGRGPDPAHVAVNL 75
 DB 266 TPPLPPVAPSPRPVQHTPTVIS 289
 RESULT 22
 ID NTC3 MOUSE STANDARD; PRT; 2318 AA.
 AC Q61982;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neurogenic locus notch homolog protein 3 precursor (Notch 3).
 GN NOTCH3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

FT REPEAT 1972 2001 ANK 5.
 FT SITE 1573 1573 CLEAVAGE (BY FURIN-LIKE PROTEASE) (BY
 FT DISULFID 43 55 SIMILARITY).
 FT DISULFID 49 66 BY SIMILARITY.
 FT DISULFID 68 77 BY SIMILARITY.
 FT DISULFID 83 94 BY SIMILARITY.
 FT DISULFID 88 107 BY SIMILARITY.
 FT DISULFID 109 118 BY SIMILARITY.
 FT DISULFID 124 135 BY SIMILARITY.
 FT DISULFID 129 145 BY SIMILARITY.
 FT DISULFID 147 156 BY SIMILARITY.
 FT DISULFID 163 175 BY SIMILARITY.
 FT DISULFID 169 184 BY SIMILARITY.
 FT DISULFID 186 195 BY SIMILARITY.
 FT DISULFID 202 213 BY SIMILARITY.
 FT DISULFID 207 223 BY SIMILARITY.
 FT DISULFID 225 234 BY SIMILARITY.
 FT DISULFID 241 252 BY SIMILARITY.
 FT DISULFID 246 261 BY SIMILARITY.
 FT DISULFID 263 272 BY SIMILARITY.
 FT DISULFID 279 292 BY SIMILARITY.
 FT DISULFID 286 301 BY SIMILARITY.
 FT DISULFID 303 312 BY SIMILARITY.
 FT DISULFID 319 330 BY SIMILARITY.
 FT DISULFID 324 339 BY SIMILARITY.
 FT DISULFID 341 350 BY SIMILARITY.
 FT DISULFID 356 367 BY SIMILARITY.
 FT DISULFID 361 378 BY SIMILARITY.
 FT DISULFID 380 389 BY SIMILARITY.
 FT DISULFID 396 409 BY SIMILARITY.
 FT DISULFID 403 418 BY SIMILARITY.
 FT DISULFID 420 429 BY SIMILARITY.
 FT DISULFID 436 447 BY SIMILARITY.
 FT DISULFID 441 456 BY SIMILARITY.
 FT DISULFID 458 467 BY SIMILARITY.
 FT DISULFID 474 485 BY SIMILARITY.
 FT DISULFID 479 494 BY SIMILARITY.
 FT DISULFID 496 505 BY SIMILARITY.
 FT DISULFID 512 523 BY SIMILARITY.
 FT DISULFID 517 532 BY SIMILARITY.

Query Match 16.7%; Score 70; DB 1; Length 2318;
 Best Local Similarity 37.7%; Pred. No. 51;
 Matches 23; Conservative 2; Mismatches 24; Indels 12; Gaps 1;

Qy 6 LPRANVPVLEMGPGPHVLSLRSNDVLSAFYSLPLAPLFTSPVSPVSGRGP 65
 Db 2162 LLNPVAVPLDWARUPPPAPPGFSEL-----LPLAPGQLLNPGAPVSPQERPP 2209

Qy 66 P 66
 Db 2210 P 2210

RESULT 23
 NTC3 HUMAN
 ID NTC3 HUMAN STANDARD; PRT: 2321 AA.
 AC Q9UM47; Q9UEB3; Q9UPL3; Q9Y6L8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neurogenic locus notch homolog protein 3 precursor (Notch 3).
 GN NOTCH3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_Taxid:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97032728; PubMed=8878478;
 RA Joutel A., Corpechot C., Ducros A., Vahedi K., Chabriat H., Mouton P.,
 RA Alamowitch S., Domenga V., Cecillion M., Marechal E., Maciazek J.,

RA Vayssiere C., Cruaud C., Cabanis E.-A., Ruchoux M.M., Weissenbach J.,
 BAch J.-P., Bousser M.-G., Tournier-Lasserre E.;
 RA "Notch3 mutations in CADASIL, a hereditary adult-onset condition
 RT causing stroke and dementia.";
 RL Nature 383:707-710(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Guel M., Artavanis-Tsakonas S.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.M.,
 RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,
 RA Phan H., Velasco N., Ganes J., Danganan L., Poundstone P.,
 RA Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andreise T.,
 RA Frankheim M., Amico-Keller G., Coefield J., Duarte S., Lucas S.,
 RA Bruce R., Thomas P., Quan G., Kronmiller B., Arellano A.,
 RA Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S.,
 RA Carrano A.V.;
 RA "Sequence analysis of an 1.5 Mb olfactory receptor (OLFR) cluster in
 RT 19p13.1";
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP VARIANTS CADASIL TYR-49; CYS-71; CYS-90; CYS-110; CYS-133; CYS-141;
 RP ARG-146; CYS-153; CYS-169; CYS-171; CYS-182; ARG-185; SER-212;
 RP GLY-222; TYR-224; CYS-258; TYR-542; CYS-558; CYS-578; CYS-728;
 RP CYS-985; CYS-1006; CYS-1031; CYS-1031 AND ARG-1261, AND VARIANTS
 RP ARG-170; LEU-496; GLN-1133; MET-1183 AND ALA-2223.
 RX MEDLINE=98049753; PubMed=9388399;
 RA Joutel A., Vahedi K., Corpechot C., Troesch A., Chabriat H.,
 RA Vayssiere C., Cruaud C., Maciazek J., Weissenbach J., Bousser M.-G.,
 RA Bach J.-P., Tournier-Lasserre E.;
 RA "Strong clustering and stereotyped nature of Notch3 mutations in
 RT CADASIL patients";
 RL Lancet 350:1511-1515(1997).
 RN [5]
 RP VARIANT CADASIL 114-GLY--PRO-120 DEL.
 RX MEDLINE=20264473; PubMed=10302807;
 RA Joutel A., Chabriat H., Vahedi K., Domenga V., Vayssiere C.,
 RA Ruchoux M.M., Lucas C., Leys D., Bousser M.-G., Tournier-Lasserre E.;
 RA "Splice site mutation causing a seven amino acid Notch3 in-frame
 RT deletion in CADASIL";
 RL Neurology 54:1874-1875(2000).
 RN [6]
 RP IDENTIFICATION OF LIGANDS.
 RX MEDLINE=99180765; PubMed=10079256;
 RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.-L.,
 RA Banks A., Leiman J., Ward D., Ish-Horowitz D., Artavanis-Tsakonas S.;
 RA "Human ligands of the Notch receptor";
 RL Am. J. Pathol. 154:785-794(1999).
 CC -I- FUNCTION: Functions as a receptor for membrane-bound ligands
 CC Jagged1, Jagged2 and Delta to regulate cell-fate determination.
 CC Upon ligand activation through the released notch intracellular
 CC domain (NICD) it forms a transcriptional activator complex with
 CC RBP-J kappa and activates genes of the enhancer of split locus.
 CC Affects the implementation of differentiation, proliferation and
 CC apoptotic programs (By similarity).
 CC -I- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-
 CC terminal fragment N(BC) which are probably linked by disulfide
 CC bonds (By similarity).
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein. Following
 CC proteolytical processing NICD is translocated to the nucleus.
 CC -I- TISSUE SPECIFICITY: Ubiquitously expressed in fetal and adult
 CC tissues.
 CC -I- PTM: Synthesized in the endoplasmic reticulum as an inactive form
 CC which is proteolytically cleaved by a furin-like convertase in the
 CC trans-Golgi network before it reaches the plasma membrane to yield
 CC an active, ligand-accessible form. Cleavage results in a C-
 CC terminal fragment N(TM) and a N-terminal fragment N(BC). Following
 CC ligand binding, it is cleaved by TNF-alpha converting enzyme
 CC (TACE) to yield a membrane-associated intermediate fragment called
 CC notch extracellular truncation (NEXT). This fragment is then
 CC cleaved by presenilin dependent gamma-secretase to release a

notch-derived peptide containing the intracellular domain (NICD)
CC from the membrane (By similarity).
CC -!- PTM: Phosphorylated (By similarity).
CC -!- DISEASE: Defects in NOTCH3 are associated with cerebral autosomal
CC dominant arteriopathy with subcortical infarcts and
CC leukoencephalopathy (CADASIL) [MIM:125310]. CADASIL causes a type
CC of stroke and dementia of which key features include recurrent
CC subcortical ischemic events and vascular dementia.
CC -!- SIMILARITY: Belongs to the NOTCH family.
CC -!- SIMILARITY: Contains 34 EGF-like domains.
CC -!- SIMILARITY: Contains 3 Lin/Notch repeats.
CC -!- SIMILARITY: Contains 5 ANK repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U97669; AAB91371.1; -;
CC DR EMBL; AF058900; AAC14346.1; -;
CC DR EMBL; AF058881; AAC14346.1; JOINED.
CC DR EMBL; AF058882; AAC14346.1; JOINED.
CC DR EMBL; AF058883; AAC14346.1; JOINED.
CC DR EMBL; AF058884; AAC14346.1; JOINED.
CC DR EMBL; AF058885; AAC14346.1; JOINED.
CC DR EMBL; AF058886; AAC14346.1; JOINED.
CC DR EMBL; AF058887; AAC14346.1; JOINED.
CC DR EMBL; AF058888; AAC14346.1; JOINED.
CC DR EMBL; AF058889; AAC14346.1; JOINED.
CC DR EMBL; AF058890; AAC14346.1; JOINED.
CC DR EMBL; AF058891; AAC14346.1; JOINED.
CC DR EMBL; AF058892; AAC14346.1; JOINED.
CC DR EMBL; AF058893; AAC14346.1; JOINED.
CC DR EMBL; AF058894; AAC14346.1; JOINED.
CC DR EMBL; AF058895; AAC14346.1; JOINED.
CC DR EMBL; AF058896; AAC14346.1; JOINED.
CC DR EMBL; AF058897; AAC14346.1; JOINED.
CC DR EMBL; AF058898; AAC14346.1; JOINED.
CC DR EMBL; AF058899; AAC14346.1; JOINED.
CC DR EMBL; AC004257; AAC04897.1; -;
CC DR EMBL; AC004663; AAC15789.1; ALT_INIT.
CC DR PIR; S78549; S78549.
CC DR HSP; P00740; IEDM.
CC
CC Genew; HGNC:7883; NOTCH3.
CC MIM; 600276; -;
CC MIM; 125310; -;
CC
CC InterPro; IPR002110; ANK.
CC InterPro; IPR000152; Asx_hydroxyl_s.
CC InterPro; IPR000742; EGF_2
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR001438; EGF_11.
CC InterPro; IPR006209; EGF_Like.
CC InterPro; IPR002049; Laminin_EGF.
CC InterPro; IPR008297; Notch.
CC InterPro; IPR000800; Notch_dom.
CC Pfam; PF00023; ank; 5
CC Pfam; PF00008; EGF_34.
CC Pfam; PF00086; notch; 3.
CC PIRSF; PIRSF002279; Notch; 1.
CC PRINTS; PR00010; EGFLOOD.
CC PRINTS; PR00011; EGF_LAMININ.
CC PRINTS; PR01452; NOTCH.
CC SMART; SM00248; ANK; 5.
CC SMART; SM00179; EGF_Ca; 19.
CC SMART; SM00004; NL; 3.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC PROSITE; PS50088; ANK_REPEAT; 4.
CC PROSITE; PS00010; ASX_HYDROXYL; 18.
CC PROSITE; PS00022; EGF_1; 33.
CC PROSITE; PS01186; EGF_2; 25.

DR PROSITE; PS50026; EGF_3; 34.
DR PROSITE; PS01187; EGF_Ca; 16.
KW Receptor; Transcription regulation; Activator; Differentiation;
KW Developmental protein; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Glycoprotein; Signal; Phosphorylation; Polymorphism;
KW Disease mutation.
FT SIGNAL 1 39
FT CHAIN 40 2321
FT CHAIN 1629 2321
FT CHAIN 1662 2321
FT DOMAIN 40 1643
FT TRANSMEM 1644 1664
FT DOMAIN 1665 2321
FT DOMAIN 40 77
FT DOMAIN 78 118
FT DOMAIN 119 156
FT DOMAIN 158 195
FT DOMAIN 197 234
FT DOMAIN 236 272
FT DOMAIN 274 312
FT DOMAIN 314 350
FT DOMAIN 351 389
FT DOMAIN 391 429
FT DOMAIN 431 467
FT DOMAIN 469 505
FT DOMAIN 507 543
FT DOMAIN 545 580
FT DOMAIN 582 618
FT DOMAIN 620 655
FT DOMAIN 657 693
FT DOMAIN 695 730
FT DOMAIN 734 770
FT DOMAIN 771 808
FT DOMAIN 810 847
FT DOMAIN 849 885
FT DOMAIN 887 922
Query Match 16.7%; Score 70; DB 1; Length 2321;
Best Local Similarity 37.7%; Pred. No. 51;
Matches 23; Conservative 2; Mismatches 24; Indels 12; Gaps 1;
QY 6 LPRPAAPVPLRMQGPAPHPVLSFLRPSWDLVSAFYSLFAPLSPTSPVSPVSGRGP 65
DB 2162 LLNPVAVPLDWARLPAPPGPSFL-----LFLAPGQLNPGTVPSPQRP 2209
QY 66 P 66
DB 2210 P 2210
RESULT 24
ID TEA3_CHICK STANDARD; PRT; 433 AA.
AC Q90701; Q90702;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Transcriptional enhancer factor TEF-5 (TEA domain family member 3)
DE (TEAD-3) (Cardiac-enriched TEA domain transcription factor 1) (TEF-
DE 1).
GN TEAD3 OR TEF5 OR TEF1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS TEF-1A AND TEF-1B).
RC TISSUE=Heart;
RX MEDLINE=96215227; PubMed=8626520;
RA Azakie A., Larkin S.B., Farrance I.K., Grenningloh G., Ordahl C.P.;

"DTEF-1, a novel member of the transcription enhancer factor-1 (TEF-1) multigene family." J. Biol. Chem. 271:8260-8265(1996).

CC -!- FUNCTION: Sequence specific M-CAT-binding factor.

CC -!- SUBCELLULAR LOCATION: Nuclear (potential).

CC -!- ALTERNATIVE PRODUCTS:

CC Name=DTEF-1A;

CC Name=DTEF-1B;

CC IsoId=Q90701-1; Sequence=Displayed;

CC IsoId=Q90701-2; Sequence=VSP 006389;

CC TISSUE SPECIFICITY: High levels in cardiac muscle, low in skeletal muscle. Intermediate levels in gizzard and lung, low levels in kidney.

CC -!- SIMILARITY: Contains 1 TEA DNA-binding domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: U46127; AAC59786.1; ALT_INIT.

CC EMBL: U46128; AAC59787.1; ALT_INIT.

CC InterPro: IPR000818; TEA/ATTS.

CC Pfam: PF01285; TEA; 1.

CC PRINTS: PR00065; TEADOMAIN.

CC SMART: SM00426; TEA; 1.

CC PROSITE: PS00554; TEA DOMAIN; 1.

CC Transcription regulation, Activator; DNA-binding; Nuclear protein; Alternative splicing.

CC DNA BIND 30 97 TEA-DOMAIN.

CC DOMAIN 146 210 PRO-RICH.

CC DOMAIN 171 433 TRANSCRIPTIONAL ACTIVATION (POTENTIAL).

CC DOMAIN 146 151 POLY-PRO.

CC VARSPLIC 87 110 KKVSSHLQVLRREISGDSSKLK -> RKQVSSSHIQVLAR

CC KKVRSSTGWHQ (in isoform DTEF-1B).

CC /FTID=VSP 006389.

CC SEQUENCE 433 AA; 48514 MW; 81077A9C4B0F245C CRC64;

Query Match 16.6%; Score 69.5; DB 1; Length 433;

Best Local Similarity 35.0%; Pred. No. 9.7;

Matches 21; Conservative 11; Mismatches 17; Indels 11; Gaps 4;

QY 4 SLLPRAAVPVPLRMQGPAPHPVLSRLPSWDVLS-----AFYSYSLAPLSP--TSVPI 55

DB 156 SAAAPRFGSGIP--GGPGSQDIKFPQAPYPIQPPMPPLASLYE-PLAPLPMAASAVPV 212

RESULT 25

PCD7_MOUSE

ID PCD7_MOUSE STANDARD; PRT; 482 AA.

AC Q8WTY1; Q8RSD9;

DT 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Programmed cell death protein 7 (ES18).

GN PDCD7.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Embryonic stem cells, and Thymus;

RX MEDLINE=99157097; PubMed=10037816;

RA Park E.J., Kim J.H., Seong R.H., Kim C.G., Park S.D., Hong S.H.;

RT "Characterization of a novel mouse cDNA, ES18, involved in apoptotic cell death of T-cells."

RL Nucleic Acids Res. 27:1524-1530(1999).

RN [2]

SEQUENCE OF 296-482 FROM N.A.

MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan E., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.B., Scheetz T.B., Brownstein M.J., Ugin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butler A.E., Schein J.E., Jones S.J.M., Skalska U., Smalusz D.B., Schnerch F., Schen J.B., Jones S.J.M., Marra M.A.;

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [3]

RP FUNCTION.

RP MEDLINE=99187471; PubMed=10087507;

RA Ritzel M.W., Yao S.Y., Ng A.M., Mackey J.R., Cass C.E., Young J.D.;

RT "Molecular cloning, functional expression and chromosomal localization of a cDNA encoding a human Na+/nucleoside cotransporter (hCNT2) selective for purine nucleosides and uridine."

RT Mol. Membr. Biol. 15:203-211(1998).

CC -!- FUNCTION: Promotes apoptosis when overexpressed.

CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).

CC -!- TISSUE SPECIFICITY: Highly expressed in testis, thymus and lymph nodes. Detected at low levels in embryonic stem cells.

CC -!- INDUCTION: Upon induction of apoptosis in embryonic stem cells by treatment with dexamethasone, staurosporine or C2-ceramide.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: AF083929; AAD20240.1; -.

DR EMBL: BC022772; AAB22772.1; ALT_INIT.

DR MGD; MGI:1859170; Pdcd7.

DR GO; GO:0006915; P:apoptosis; IDA.

KW Apoptosis; Nuclear protein; Coiled coil.

FT DOMAIN 4 129 PRO-RICH.

FT DOMAIN 139 275 ARG-RICH.

FT DOMAIN 233 408 COILED COIL (POTENTIAL).

SQ SEQUENCE 482 AA; 54357 MW; 622E986F177C00A1 CRC64;

Query Match 16.6%; Score 69.5; DB 1; Length 482;

Best Local Similarity 34.2%; Pred. No. 11;

Matches 27; Conservative 4; Mismatches 25; Indels 23; Gaps 5;

QY 6 LPRPAAVPVPLRMQGPAPHPVLS--FLRPSWDVLS-----SAFYSYSLAPLSP 49

DB 32 LPSP-AFPPPLPQPGP-FCGASAPFLQPLAPLPAPFAEASRGCGGAGFPVPPPLP 89

QY 50 PTVSPISPVSVGRGDDPA 68

DB 90 PFPQCRPF-----FQPD 103

RESULT 26

ENAH_MOUSE

ID ENAH_MOUSE STANDARD; PRT; 802 AA.

AC Q03173; P70430; P70431; P70432; P70433;

DT 01-OCT-1996 (Rel. 34, Created)

"DTEF-1, a novel member of the transcription enhancer factor-1 (TEF-1) multigene family." J. Biol. Chem. 271:8260-8265(1996).

CC -!- FUNCTION: Sequence specific M-CAT-binding factor.

CC -!- SUBCELLULAR LOCATION: Nuclear (potential).

CC -!- ALTERNATIVE PRODUCTS:

CC Name=DTEF-1A;

CC Name=DTEF-1B;

CC IsoId=Q90701-1; Sequence=Displayed;

CC IsoId=Q90701-2; Sequence=VSP 006389;

CC TISSUE SPECIFICITY: High levels in cardiac muscle, low in skeletal muscle. Intermediate levels in gizzard and lung, low levels in kidney.

CC -!- SIMILARITY: Contains 1 TEA DNA-binding domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: U46127; AAC59786.1; ALT_INIT.

CC EMBL: U46128; AAC59787.1; ALT_INIT.

CC InterPro: IPR000818; TEA/ATTS.

CC Pfam: PF01285; TEA; 1.

CC PRINTS: PR00065; TEADOMAIN.

CC SMART: SM00426; TEA; 1.

CC PROSITE: PS00554; TEA DOMAIN; 1.

CC Transcription regulation, Activator; DNA-binding; Nuclear protein; Alternative splicing.

CC DNA BIND 30 97 TEA-DOMAIN.

CC DOMAIN 146 210 PRO-RICH.

CC DOMAIN 171 433 TRANSCRIPTIONAL ACTIVATION (POTENTIAL).

CC DOMAIN 146 151 POLY-PRO.

CC VARSPLIC 87 110 KKVSSHLQVLRREISGDSSKLK -> RKQVSSSHIQVLAR

CC KKVRSSTGWHQ (in isoform DTEF-1B).

CC /FTID=VSP 006389.

CC SEQUENCE 433 AA; 48514 MW; 81077A9C4B0F245C CRC64;

Query Match 16.6%; Score 69.5; DB 1; Length 433;

Best Local Similarity 35.0%; Pred. No. 9.7;

Matches 21; Conservative 11; Mismatches 17; Indels 11; Gaps 4;

QY 4 SLLPRAAVPVPLRMQGPAPHPVLSRLPSWDVLS-----AFYSYSLAPLSP--TSVPI 55

DB 156 SAAAPRFGSGIP--GGPGSQDIKFPQAPYPIQPPMPPLASLYE-PLAPLPMAASAVPV 212

RESULT 25

PCD7_MOUSE

ID PCD7_MOUSE STANDARD; PRT; 482 AA.

AC Q8WTY1; Q8RSD9;

DT 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Programmed cell death protein 7 (ES18).

GN PDCD7.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Embryonic stem cells, and Thymus;

RX MEDLINE=99157097; PubMed=10037816;

RA Park E.J., Kim J.H., Seong R.H., Kim C.G., Park S.D., Hong S.H.;

RT "Characterization of a novel mouse cDNA, ES18, involved in apoptotic cell death of T-cells."

RL Nucleic Acids Res. 27:1524-1530(1999).

RN [2]

SEQUENCE OF 296-482 FROM N.A.

MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan E., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.B., Scheetz T.B., Brownstein M.J., Ugin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butler A.E., Schein J.E., Jones S.J.M., Skalska U., Smalusz D.B., Schnerch F., Schen J.B., Jones S.J.M., Marra M.A.;

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [3]

RP FUNCTION.

RP MEDLINE=99187471; PubMed=10087507;

RA Ritzel M.W., Yao S.Y., Ng A.M., Mackey J.R., Cass C.E., Young J.D.;

RT "Molecular cloning, functional expression and chromosomal localization of a cDNA encoding a human Na+/nucleoside cotransporter (hCNT2) selective for purine nucleosides and uridine."

RT Mol. Membr. Biol. 15:203-211(1998).

CC -!- FUNCTION: Promotes apoptosis when overexpressed.

CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).

CC -!- TISSUE SPECIFICITY: Highly expressed in testis, thymus and lymph nodes. Detected at low levels in embryonic stem cells.

CC -!- INDUCTION: Upon induction of apoptosis in embryonic stem cells by treatment with dexamethasone, staurosporine or C2-ceramide.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: AF083929; AAD20240.1; -.

DR EMBL: BC022772; AAB22772.1; ALT_INIT.

DR MGD; MGI:1859170; Pdcd7.

DR GO; GO:0006915; P:apoptosis; IDA.

KW Apoptosis; Nuclear protein; Coiled coil.

FT DOMAIN 4 129 PRO-RICH.

FT DOMAIN 139 275 ARG-RICH.

FT DOMAIN 233 408 COILED COIL (POTENTIAL).

SQ SEQUENCE 482 AA; 54357 MW; 622E986F177C00A1 CRC64;

Query Match 16.6%; Score 69.5; DB 1; Length 482;

Best Local Similarity 34.2%; Pred. No. 11;

Matches 27; Conservative 4; Mismatches 25; Indels 23; Gaps 5;

QY 6 LPRPAAVPVPLRMQGPAPHPVLS--FLRPSWDVLS-----SAFYSYSLAPLSP 49

DB 32 LPSP-AFPPPLPQPGP-FCGASAPFLQPLAPLPAPFAEASRGCGGAGFPVPPPLP 89

QY 50 PTVSPISPVSVGRGDDPA 68

DB 90 PFPQCRPF-----FQPD 103

RESULT 26

ENAH_MOUSE

ID ENAH_MOUSE STANDARD; PRT; 802 AA.

AC Q03173; P70430; P70431; P70432; P70433;

DT 01-OCT-1996 (Rel. 34, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Enabled protein homolog (NPC derived proline-rich protein 1) (NDPP-1).
GN ENAH OR MENA OR NDPP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=93041923; PubMed=1420303;
RA Sazuka T., Tomooka Y., Kathju S., Ikawa Y., Noda M., Kumar S.;
RT Identification of a developmentally regulated gene in the mouse
central nervous system which encodes a novel proline rich protein.";
RL Biochim. Biophys. Acta 1132:240-248(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 2; 3; 4 AND 5), FUNCTION, AND SUBCELLULAR
RP LOCATION.
RC TISSUE=Brain;
RX MEDLINE=97015079; PubMed=8861907;
RA Gertler F.B., Niebuhr K., Reinhard M., Wehland J., Soriano P.;
RT "Mena, a relative of VASP and Drosophila Enabled, is implicated in the
control of microfilament dynamics.";
RL Cell 87:227-239(1996).
RN [3]
RP FUNCTION, AND SUBUNIT.
RX MEDLINE=99166867; PubMed=10069337;
RA Lanier L.M., Gates M.A., Witke W., Menzies A.S., Wehman A.M.,
RA Macklis J.D., Kwiatkowski D., Soriano P., Gertler F.B.;
RT "Mena is required for neurulation and commissure formation.";
RL Neuron 22:313-325(1999).
CC -!- FUNCTION: May be involved in microfilament assembly and cell
motility. Induces the formation of F-actin rich outgrowths in
fibroblasts. Required for neurulation and commissure formation.
CC -!- SUBUNIT: Binds profilin.
CC -!- SUBCELLULAR LOCATION: Localized to focal adhesions and, to a
lesser extent, leading edges and stress fibers.
CC -!- ALTERNATIVE PRODUCTS.
CC Name=Alternative splicing; Named isoforms=5;
CC Name=5; Synonyms=Mena+;
CC IsoID=Q03173-1; Sequences=Displayed;
CC Name=1;
CC IsoID=Q03173-2; Sequences=VSP_007255, VSP_007256;
CC Name=2; Synonyms=Mena;
CC IsoID=Q03173-3; Sequences=VSP_007259, VSP_007260;
CC Name=3; Synonyms=Mena+;
CC IsoID=Q03173-4; Sequences=VSP_007259;
CC Name=4; Synonyms=Mena+;
CC IsoID=Q03173-5; Sequences=VSP_007257, VSP_007258;
CC -!- TISSUE SPECIFICITY: In the heart and testis and less so in the
lung, skeletal muscle, kidney, pancreas and brain.
CC -!- SIMILARITY: Contains 1 WH1 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC
CC EMBL: D10727; BAA01570.1; -
CC EMBL: U72520; AAC52863.1; -
CC EMBL: U72521; AAC52864.1; -
CC EMBL: U72522; AAC52865.1; -
CC EMBL: U72523; AAC52866.1; -
CC PIR: S27200; S27200.
CC MGD: MG1:108360; Enah.
CC GO: GO:0005925; C:focal adhesion; IDA.
CC GO: GO:0005522; F:profilin binding; IDA.
CC GO: GO:0017124; F:SH3-domain binding; IDA.
CC GO: GO:0007015; P:actin filament organization; IDA.

DR GO: GO:0007411; P:axon guidance; IMP.
DR GO: GO:0006928; P:cell motility; IDA.
DR GO: GO:0001679; P:neurulation; IDA.
DR InterPro: IPR000697; EVH1.
DR InterPro: IPR001960; WH1.
DR SMART: SM00461; WH1; 1.
KW Developmental protein; Neurogenesis; Alternative splicing.
FT DOMAIN 1 108
FT DOMAIN 442 464 POLY-PRO.
FT DOMAIN 542 552 POLY-PRO.
FT DOMAIN 562 574 POLY-PRO.
FT DOMAIN 578 589 POLY-PRO.
FT DOMAIN 593 605 POLY-PRO.
FT VARSPLIC 1 412 Missing (in isoform 1).
FT VARSPLIC 535 631 Missing (in isoform 1).
FT VARSPLIC 535 631 Missing (in isoform 1).
FT VARSPLIC 117 131 Missing (in isoform 4).
FT VARSPLIC 132 135 CIPC -> VFYL (in isoform 4).
FT VARSPLIC 117 135 Missing (in isoform 2 and isoform 3).
FT VARSPLIC 259 500 Missing (in isoform 2).
FT VARSPLIC 802 AA; 85844 MW; 592BB975EE20F77F CRC64;
SQ SEQUENCE 802 AA; 85844 MW; 592BB975EE20F77F CRC64;
Query Match 16.6%; Score 69.5; DB 1; Length 802;
Best Local Similarity 33.3%; Pred. No. 19;
Matches 28; Conservative 5; Mismatches 28; Indels 23; Gaps 5;
QY 3 HSLVPRFAAIVPVRMOPGPAHVPVLSFLRPS-----WDLVSAFYSPLAPLSPTS 52
DB 334 HVLPHVATVPRLNKNRSPSPVNT---PSSQPPAAKSCAWP-TSNFSLPSP----- 384
QY 53 VPI---SPVSVGRGPDPAHVAVN 73
DB 385 -PIMISSPPGKATGPRVLPVVCVS 407
RESULT 27
HCN2 HUMAN
ID HCN2 HUMAN STANDARD; PRT: 889 AA.
AC Q9UL51; Q60742; Q60743; Q75267; Q9UBS2;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Potassium/sodium hyperpolarization-activated cyclic nucleotide-gated
channel 2 (Brain cyclic nucleotide gated channel 2) (BCNG-2).
GN HCN2 OR BCNG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9453744; PubMed=10524219;
RA Vaccari T., Moroni A., Rocchi M., Gorza L., Bianchi M.E., Beltrame M.,
RA DiFrancesco D.;
RT "The human gene coding for HCN2, a pacemaker channel of the heart.";
RL Biochim. Biophys. Acta 1446:419-425(1999).
RN [2]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE=Heart.
RX MEDLINE=99246258; PubMed=10228147;
RA Ludwig A., Zong X., Stieber J., Füllin R., Hofmann F., Biel M.;
RT "Two pacemaker channels from human heart with profoundly different
activation kinetics.";

EMBO J. 18:2323-2329(1999).
[3]
SEQUENCE OF 1-528 FROM N.A.
Lamerdin J.E., McCready P.M., Dunn J., Skowronski E., Adamson A.W.,
Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stiliwagen S.,
Phan H., Velasco N., Do L., Regala W., Terry A., Gaines J.,
Dangman L., Poundstone P., Christensen M., Georgescu A., Brower A.,
Avila J., Liu S., Attix C., Andreise T., Trankheim M.,
Amico-Keller G., Coesfield J., Duarte S., Lucas S., Bruce R.,
Thomas P., Quan S., Krommiller B., Arellano A., Montgomery M., Ow D.,
Carraro A.V.,
"Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a
serine protease gene cluster."
Submitted (JAN-2001) to the EMBL/GenBank/DDJ databases.
[4]
SEQUENCE OF 147-743 FROM N.A., AND TISSUE SPECIFICITY.
TISSUE=Brain;
MEDLINE=98292171; PubMed=9630217;
Santoro B., Liu D.T., Yao H., Bartsch D., Kandel E.R.,
Siegelbaum S.A., Tibbs G.R.;
"Identification of a gene encoding a hyperpolarization-activated
pacemaker channel of brain."
Cell 93:717-729(1998).
-!- FUNCTION: Hyperpolarization-activated ion channel exhibiting weak
selectivity for potassium over sodium ions. Contributes to the
native pacemaker currents in heart (if) and in neurons (ih).
Produces a large instantaneous current. Activated by cAMP.
Modulated by intracellular chloride ions and pH; acidic pH shifts
the activation to more negative voltages (by similarity).
-!- SUBUNIT: The potassium channel is probably composed of a homo- or
heterotetrameric complex of pore-forming subunits. Heteromultimer
with HCN1. Interacts with KCNE2 (by similarity).
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- TISSUE SPECIFICITY: Highly expressed throughout the brain.
Detected at low levels in heart.
-!- DOMAIN: The segment S4 is probably the voltage-sensor and is
characterized by a series of positively charged amino acids at
every third position.
-!- MISCELLANEOUS: Inhibited by extracellular cesium ions.
-!- SIMILARITY: Belongs to the potassium channel family. HCN
subfamily.
-!- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF065164; AAC28444.2; -;
EMBL; AF012582; CAB42602.1; -;
EMBL; AJ133727; CAB42630.1; -;
EMBL; AJ133728; CAB42630.1; JOINED.
EMBL; AJ133729; CAB42630.1; JOINED.
EMBL; AJ133730; CAB42630.1; JOINED.
EMBL; AJ133731; CAB42630.1; JOINED.
EMBL; AJ133732; CAB42630.1; JOINED.
EMBL; AJ133733; CAB42630.1; JOINED.
EMBL; AJ133734; CAB42630.1; JOINED.
EMBL; AC005559; AAC33280.2; -;
EMBL; AF064877; AAC33760.1; -;
EMBL; HGNC:4846; HCN2.
MIM: 602781; -;
InterPro: IPR000595; cNMP binding.
InterPro: IPR005821; Ion Trans.
InterPro: IPR001622; K+channel pore.
InterPro: IPR005820; M+channel nlg.
Pfam: PF000027; cNMP binding; 1.
Pfam: PF00520; ion Trans; 1.
SMART; SM00100; cNMP; 1.

DR PROSITE; PS00888; CNMP_BINDING_1; 1.
DR PROSITE; PS00889; CNMP_BINDING_2; FALSE_NEG.
DR PROSITE; PS00042; CNMP_BINDING_3; 1.
KW Transport; Ion transport; Ionic channel; Voltage-gated channel;
KW Potassium channel; Potassium; Potassium transport; Sodium transport;
KW CAMP; CAMP-binding; Transmembrane; Glycoprotein; Sodium channel.
FT DOMAIN 1 215 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 216 236 SEGMENT S1 (POTENTIAL).
FT TRANSMEM 241 261 SEGMENT S2 (POTENTIAL).
FT DOMAIN 262 288 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 289 309 SEGMENT S3 (POTENTIAL).
FT TRANSMEM 318 338 SEGMENT S4 (POTENTIAL).
FT DOMAIN 339 369 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 370 390 SEGMENT S5 (POTENTIAL).
FT TRANSMEM 414 435 SEGMENT H5 (POT-FORMING) (POTENTIAL).
FT TRANSMEM 441 461 SEGMENT S6 (POTENTIAL).
FT DOMAIN 462 489 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 10 140 PRO-RICH.
FT DOMAIN 158 209 INVOLVED IN SUBUNIT ASSEMBLY (BY
SIMILARITY).
FT NP_BIND 544 661 CAMP.
FT DOMAIN 715 861 PRO-RICH.
FT CARBOHYD 407 407 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 17 20 TPAP -> SPTT (IN REF. 1).
FT CONFLICT 29 29 A -> R (IN REF. 1).
FT CONFLICT 32 32 Q -> K (IN REF. 1).
FT CONFLICT 294 294 D -> V (IN REF. 4).
FT CONFLICT 713 713 L -> F (IN REF. 4).
FT CONFLICT 849 849 R -> G (IN REF. 1).
SQ SEQUENCE 889 AA; 97049 MW; 5A263E1D06C2A47D CRC64;
Query Match 16.6%; Score 69.5; DS 1; Length 889;
Best Local Similarity 11.3%; Pred. No. 21;
Matches 21; Conservative 8; Mismatches 35; Indels 3; Gaps 1;
Qy 1 GTSHLLPRPAAYVPLRMQPG---PAHVLFLRPSMDLVSAFYSLPLAELSPTSVISP 57
Db 753 GSPRLVRPPPGAPAAASPGPPPPAPAPRPTSPYGLGPAALGAPALPARR 812
Qy 58 VSVGRGP 64
Db 813 LSRASRP 819
RESULT 28
RIN3 MOUSE
ID_RIN3_MOUSE STANDARD; PRT; 980 AA.
AC PS9729;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ras and Rab interactor 3 (Ras interaction/interference protein 3).
GN RIN3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory bulb;
EX MEDLINE=22354683; PubMed=12456851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kogayama A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,

RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
CC -!- FUNCTION: Potential Ras effector protein. May function as a
CC guanine nucleotide exchange (GEF), by exchanging bound GDP for
CC free GTP (By similarity).
CC -!- SUBUNIT: Interacts with RAB5B (By similarity).
CC -!- SIMILARITY: Belongs to the RIN (Ras interaction/interference)
CC family.
CC -!- SIMILARITY: Contains 1 Ras-associating domain.
CC -!- SIMILARITY: Contains 1 SH2 domain.
CC -!- SIMILARITY: Contains 1 VPS9 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AK032205; BAC27757.1; .
CC InterPro; IPR000159; RA domain.
CC InterPro; IPR000980; SH2.
CC InterPro; IPR003123; VPS9.
CC Pfam; PF00788; RA; 1.
CC Pfam; PF02204; VPS9; 1.
CC SMART; SM00314; RA; 1.
CC SMART; SM00167; VPS9; 1.
CC PROSITE; PS02000; RA; 1.
CC PROSITE; PS00001; SH2; FALSE NEG.
CC KW GTPase activation; SH2 domain.
CC FT DOMAIN 63 158 SH2.
CC FT DOMAIN 735 839 VPS9.
CC FT DOMAIN 865 962 RAS-ASSOCIATING.
CC FT DOMAIN 242 380 PRO-RICH.
CC FT DOMAIN 719 722 POLY-THR.
CC SEQUENCE 980 AA; 107275 MW; 55C562418961BE24 CRC64;
Query Match 16.6%; Score 69.5; DB 1; Length 980;
Best Local Similarity 32.4%; Pred. No. 23;
Matches 24; Conservative 7; Mismatches 32; Indels 11; Gaps 3;
QY 7 PRRAVVPVPLRMQP--GPAHVLVFLRPSNDLVSAFYSLPLAPLSPTS--VPISPVSVGR 62
Db 273 PRPPPPPPPLPTVPETGPARPLAPPVPAG-----PLPNSPLTPTSLAPHPGPPGH 325
QY 63 GPDPAHVAVNLRS 76
Db 326 SNQPPMKACESLPR 339
RESULT 29
ID_CDX1_HUMAN STANDARD; PRT; 265 AA.
AC P47902; Q9NYK8;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Homeobox protein CDX-1 {Caudal-type homeobox protein 1}.
GN CDX1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RX MEDLINE=96015048; PubMed=8530027;
RA Bonner C.A., Iofus S.K., Wasmuth J.J.;
RT "Isolation, characterization, and precise physical localization of
RT human CDX1, a caudal-type homeobox gene.";
RL Genomics 28:206-211 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon carcinoma;
RX MEDLINE=97188282; PubMed=9036867;
RA Mallo G.V., Rechste H., Frigerio J.M., Rocha D., Zweibaum A.,
RA Lacasa M., Jordan B.R., Dusetti N.J., Dagnon J.C., Iovanna J.L.;
RT "Molecular cloning, sequencing and expression of the mRNA encoding
RT human Cdx1 and Cdx2 homeobox. Down-regulation of Cdx1 and Cdx2 mRNA
RT expression during colorectal carcinogenesis.";
RL Int. J. Cancer 74:35-44 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC Malakooti J.;
RT "Molecular cloning and sequencing of the human CDX1 homeobox gene.";
Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Could play a role in the terminal differentiation of the
CC intestine.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Intestinal epithelium.
CC -!- SIMILARITY: Belongs to the Caudal homeobox family.
CC -!- SIMILARITY: Contains 1 homeobox domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U1360; AAA80284.1; .
CC EMBL; U15212; AAC50237.1; .
CC EMBL; U51095; AAB40602.1; .
CC EMBL; AF239666; AAF61234.1; .
CC HSSP; P14653; I872.
CC TRANSFAC; T03248; .
CC Genew; HGNC:1805; CDX1.
CC MIM; 600746; .
CC GO; GO:0007397; P:histogenesis and organogenesis; TAS.
CC InterPro; IPR006820; Caudal act.
CC InterPro; IPR001356; Homeobox.
CC InterPro; IPR000047; HTH lambdarepressor.
CC Pfam; PF04731; Caudal_act; 1.
CC Pfam; PF00046; Homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC PRINTS; PR00031; HTHREPRESSR.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT DOMAIN 143 149 POLY-GLY.
FT DNA_BIND 154 213 HOMEBOX.
FT DOMAIN 217 221 POLY-GLN.
FT CONFLICT 28 29 AA -> QA (IN REF. 3).
SQ SEQUENCE 265 AA; 28124 MW; DDEA69E956DEB42C CRC64;
Query Match 16.5%; Score 69; DB 1; Length 265;
Best Local Similarity 31.0%; Pred. No. 6.5;

Matches 22; Conservative 6; Mismatches 25; Indels 18; Gaps 3;
QY 9 PAAVVPV-----LRMCPGPAHV---LSFLSPSWDLVSFAFYSLEPLAPLSPTSVIS 56
Db 33 PAPPAPQYDFSSYSHVERAPAPPTAGCAFPFPAKDDWAAAYGFGAAAPAA-----S 86
QY 57 PVSVGRGPPDP 67
Db 87 PASLAFGPPDP 97

Query Match 16.4%; Score 68.5; DB 1; Length 886;
Best Local Similarity 32.9%; Pred.No.26;
Matches 26; Conservative 8; Mismatches 34; Indels 11; Gaps 5;
QY 1 GTH-----SLLPRPAAVFPVPLRMCPGPAHVLSFL-RESMD-----LVSAFYSLPLAPLSPT 51
Db 699 GPHDLDTGLLPTTPECTPLPKRLPTP-HPHAHALGSPANDHSHALLSASASTSLLLLAPA 757
QY 52 SVPISPVSVGRGPPDAHV 70
Db 758 RASEQP-QVPAERGPESRL 775

Search completed: July 4, 2004, 04:19:18
Job time : 7.80723 secs

RESULT 30
SM6B MOUSE
ID SM6B MOUSE STANDARD; PRT; 886 AA.
AC O54951;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Semaphorin 6B precursor (Semaphorin VIB) (Sema VIB) (Semaphorin N)
DE (Sema N).
GN SEMA6B OR SEMAN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98027184; PubMed=9361278;
RA Eckhardt F., Behar O., Calautti E., Yonezawa K., Nishimoto I.,
RA Fishman M.C.;
RL Mol. Cell. Neurosci. 9:409-419(1997).
CC -!- SUBUNIT: HOMODIMER. BINDS SPECIFICALLY THE SH3 DOMAIN OF THE
CC PROTOONCOGENE C-SRC.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: During development it is expressed in
CC subregions of the nervous system and is particularly prominent in
CC muscle. In adulthood, it is expressed ubiquitously.
CC -!- SIMILARITY: Belongs to the semaphorin family.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF036585; AAC00493.1; -
CC MGD; MGI:1202889; Sema6B.
CC InterPro; IPR003659; Plexin-like.
CC InterPro; IPR001627; Sema.
CC Pfam; PF01403; Sema; 1.
CC SMART; SM00423; PSI; 1.
CC SMART; SM00630; Sema; 1.
CC Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
CC Developmental protein.
CC
CC SIGNAL 1 26
CC CHAIN 27 886
CC DOMAIN 27 605
CC TRANSMEM 606 626
CC DOMAIN 627 886
CC DOMAIN 239 549
CC DOMAIN 751 754
CC POLY-LEU.
CC CARBOHYD 75 75
CC CARBOHYD 156 156
CC CARBOHYD 292 292
CC CARBOHYD 387 387
CC CARBOHYD 442 442
CC CARBOHYD 463 463
CC N-TERMINAL 886 AA; 95466 MW; E5F56D125CDA574D CRC64;
SQ SEQUENCE

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	406	97.1	419	4	Q9UK79	Q9UK79 homo sapien
2	99.5	23.8	309	12	Q8UZB4	Q8UZB4 grapevine f
3	88	21.1	122	12	Q8AZM4	Q8AZM4 hepatitis e
4	87	20.8	122	12	Q8BS06	Q8BS06 swine hepat
5	86	20.6	369	11	Q8BN79	Q8BN79 mus musculu
6	86	20.6	369	11	Q7TNS8	Q7TNS8 mus musculu
7	85	20.3	327	11	Q99JK6	Q99JK6 mus musculu
8	84.5	20.2	434	5	Q9GU15	Q9GU15 hepatitis p
9	84	20.1	119	12	Q80IR8	Q80IR8 hepatitis e
10	84	20.1	995	11	Q35615	Q35615 mus musculu
11	82.5	19.7	316	10	Q8R2V1	Q8R2V1 oryza sativ
12	82	18.6	122	12	Q8V730	Q8V730 swine hepat
13	82	18.6	122	12	Q91115	Q91115 hepatitis e
14	81.5	18.5	816	11	Q70474	Q70474 rattus norv
15	81	18.1	200	10	Q82761	Q82761 arabidopsis
16	81	19.4	1006	10	Q9LMQ1	Q9LMQ1 arabidopsis

```
90 71 17.0 189 12 Q91QR2 Q91qr2 soybean dwa
ALIGNMENTS
RESULT 1
Q9UK79
ID Q9UK79 PRELIMINARY; PRT; 419 AA.
AC Q9UK79;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Herstatin.
GN HER-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99415951; PubMed=10485918;
RA Doherty J.K., Bond C., Jardim A., Adelman J.P., Clinton G.M.;
RT "The HER-2/neu receptor tyrosine kinase gene encodes a secreted
RT autoinhibitor.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10869-10874 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Doherty J.K., Clinton G.M., Adelman J.P., Evans A.J., Henner W.D.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF177761; AAD56009.2; -.
DR GO; GO:0005576; C:extracellular; IDA.
DR GO; GO:0004715; P:non-membrane spanning protein tyrosine kin. . ; IDA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . ; IDA.
DR InterPro; IPR000494; EGFR_L domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow_fac_recep.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L domain; 1.
DR SMART; SM00261; FU; 1.
DR SEQUENCE 419 AA; 45472 MW; FEC1BE347E2D030C CRC64;
Query Match 97.1%; Score 406; DB 4; Length 419;
Best Local Similarity 97.5%; Pred. No. 3.3e-35;
Matches 77; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GTHSLPRPAAPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVPSV 60
DB 341 GTHSLPRPAAPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPVPSV 400
QY 61 GRGPDPAHVAVNLSYEG 79
DB 401 GRGPDPAHVAVNLSYEG 419
RESULT 2
Q8U2B4
ID Q8U2B4 PRELIMINARY; PRT; 309 AA.
AC Q8U2B4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Grapevine fleck virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage.
OX NCBI_TaxID=103722;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MT48;
RA Sabanadzovic S.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=MT48;
RX MEDLINE=21351056; PubMed=11458008;
RA Sabanadzovic S.; Abou Ghanem-Sabanadzovic N., Saidarelli P.,
RA Martelli G.P.;
RT "Complete nucleotide sequence and genome organisation of Grapevine
RT fleck virus.";
RL J. Gen. Virol. 82:2009-2015 (2001).
DR EMBL; AJ309022; CAC84402.1; -.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR InterPro; IPR003882; Pistil_extensin.
DR PRINTS; PR01218; PSTLEXTENSIN.
KW Hypothetical protein.
SQ SEQUENCE 309 AA; 31480 MW; E40FG8A7F54737B2 CRC64;
Query Match 23.8%; Score 99.5; DB 12; Length 309;
Best Local Similarity 39.0%; Pred. No. 0.011;
Matches 30; Conservative 8; Mismatches 20; Indels 19; Gaps 5;
QY 6 LPRPAAPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPV 52
DB 176 VPRSAAPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPV 229
QY 53 VPISPVSVGRCPDPAH 69
DB 230 PELSPLSPLHAPAPPPH 246
RESULT 3
Q8AZM4
ID Q8AZM4 PRELIMINARY; PRT; 122 AA.
AC Q8AZM4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Genomic RNA, complete genome, isolate:HE-JA10.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HE-JA10;
RA Okamoto H.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HE-JA10;
RX MEDLINE=22447293; PubMed=12560575;
RA Tokita H., Harada H., Gotanda Y., Takahashi M., Nishizawa T.,
RA Okamoto H.;
RT "Molecular and serological characterization of sporadic acute
RT hepatitis E in a Japanese patient infected with a genotype III
RL hepatitis E virus in 1993.";
RL J. Gen. Virol. 84:421-427 (2003).
DR EMBL; AB089824; BAC44898.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR InterPro; IPR003384; HSV_ORF2_1.
DR Pfam; PF02444; HSV_ORF2_1.
DR SEQUENCE 122 AA; 12369 MW; 368D92E21F85A0B0 CRC64;
Query Match 21.1%; Score 98; DB 12; Length 122;
Best Local Similarity 38.3%; Pred. No. 0.071;
Matches 23; Conservative 10; Mismatches 21; Indels 6; Gaps 3;
QY 1 GTHSLPRPAAPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSPV 57
DB 55 GVTGLIPSPS--PSPIFIQPTFS--PPMSFHHFGLALDSRPAPLAPLGVTSAPAPLP 111
RESULT 4
Q8BB06
```

```

ID Q8B06 PRELIMINARY; PRT; 122 AA.
AC Q8B06;
DT 01-MAR-2003 (TrEMBLrel. 23, Created);
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update);
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update);
DE Cytokeleton-associated phosphoprotein.
OS Swine hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=63421;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=Arkell;
RX MEDLINE=22297120; PubMed=12409369;
RA Pei Y., Yoo D.;
RT "Genetic Characterization and Sequence Heterogeneity of a Canadian
RT Isolate of Swine Hepatitis E Virus.";
RL J. Clin. Microbiol. 40:4021-4029(2002).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN=Arkell;
RA Yoo D., Pei Y.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY115488; AM92183.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2; 1.
SQ SEQUENCE 122 AA; 12369 MW; F06A8DFC599442D CRC64;

Query Match 20.8%; Score 87; DB 12; Length 122;
Best Local Similarity 39.3%; Pred. No. 0.091;
Matches 24; Conservative 10; Mismatches 21; Indels 6; Gaps 3;

QY 1 GTHSLLPRAAVVPVPLRMQGPAPHPVLSFLRPSMDLVSAFYSLPLAPLSPTSPVPSV 57
DB 55 GVTGLSPS--PSPIQFTPS--PPMSFHPGLEALDRAPLGLVGFSPAPPLSP 111
QY 58 V 58
DB 112 V 112

RESULT 5
Q8B79 ID Q8B79 PRELIMINARY; PRT; 369 AA.
AC Q8B79;
DT 01-MAR-2003 (TrEMBLrel. 23, Created);
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update);
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update);
DE Hypothetical protein.
GN E130012A19RIK OR AA409164.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eye;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK087411; BAC39864.1; -;
DR MGD; MGI:2143991; E130012A19RIK.
KW Hypothetical protein.
SQ SEQUENCE 369 AA; 38009 MW; DFE16A97B852FBCF CRC64;

Query Match 20.6%; Score 86; DB 11; Length 369;
Best Local Similarity 42.4%; Pred. No. 0.38;
Matches 28; Conservative 4; Mismatches 28; Indels 6; Gaps 4;

QY 1 GTHSLLPRAAVVPVPLRMQGPAPHPVLSFLRPSMDLVSAFYSLPLAPLSPTSPVPSV 60
DB 130 GAGSCAP-PGRCAPAPLRQDSPTNPAPSPPPARGLDAA--SSP--PLEPGSPPPSP-PA 183
QY 61 GRGPD 66
DB 184 GLSPEP 189

RESULT 6
Q7TNS8 ID Q7TNS8 PRELIMINARY; PRT; 369 AA.
AC Q7TNS8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created);
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update);
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update);
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Kleusner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Locuelli N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Boesak S.A., McEvan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC055770; AAH55770.1; -;
KW Hypothetical protein.
SQ SEQUENCE 369 AA; 38095 MW; C93AC0213E4E97BF CRC64;

Query Match 20.6%; Score 86; DB 11; Length 369;
Best Local Similarity 42.4%; Pred. No. 0.38;
Matches 28; Conservative 4; Mismatches 28; Indels 6; Gaps 4;

QY 1 GTHSLLPRAAVVPVPLRMQGPAPHPVLSFLRPSMDLVSAFYSLPLAPLSPTSPVPSV 60
DB 130 GAGSCAP-PGRCAPAPLRQDSPTNPAPSPPPARGLDAA--SSP--PLEPGSPPPSP-PA 183
QY 61 GRGPD 66
DB 184 GLSPEP 189

RESULT 7
Q99JK6 ID Q99JK6 PRELIMINARY; PRT; 327 AA.
AC Q99JK6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created);

```



```
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
KW EMBL; BC006054; AAH06054.1; -.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 327 AA; 33661 MW; 27917F16D583E774 CRC64;

Query Match 20.3%; Score 85; DB 11; Length 327;
Best Local Similarity 42.4%; Pred. No. 0.42;
Matches 28; Conservative 4; Mismatches 28; Indels 6; Gaps 4;

QY 1 GTHSLPRPAAPVPLRMQPGPAHPVLSFLRPSMDLVSFAFYSLLPLAPLSPTSPVPSV 60
DB 86 GAGSCSP-PCRCPAPLRQDSNTNPAWSPFRPARGLDA--SSP--PLEPGSPFPSP-PA 141
QY 61 GEGPDP 66
DB 142 GLSPEP 147

RESULT 8
Q9GU15 PRELIMINARY; PRT; 434 AA.
AC Q9GU15;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-NAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Theileria parva.
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileriidae;
OC Theileria.
OX NCBI_TaxID=5875;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20520968; PubMed=11071289;
RA Bishop R., Gobright E., Nene V., Morzaria S., Musoke A., Sohanpal B.;
RT "Polymorphic open reading frames encoding secretory proteins are
RT located less than 3 kilobases from Theileria parva telomeres.";
RL Mol. Biochem. Parasitol. 110:359-371(2000).
DR EMBL; AF225701; AAG28022.1; -.
DR InterPro; IPR002965; P rich extensin.
DR PRINTS; PR01217; PRICHEXTNSN.
KW Hypothetical protein.
SQ SEQUENCE 434 AA; 49670 MW; 0F46586F4021A22F CRC64;

Query Match 20.2%; Score 84.5; DB 5; Length 434;
Best Local Similarity 31.0%; Pred. No. 0.65;
Matches 26; Conservative 11; Mismatches 34; Indels 13; Gaps 4;

QY 4 SLLPRPAAPVPLRMQPGPAHPVLSFLRPSMDLVSFAFYSLLPLAPLSPTSPVPI 55
DB 159 ALYPRPTQIPQLPPLQPGHFAFI-----PRYPQVSGYGFYHYPARPPSPVQPIPPPS 213
QY 56 SPVSVGRGDPDPAHVAVNLSRYEG 79
DB 214 THYVPTQPPQPPQPPQYQYVG 237

RESULT 9
Q80IR8 PRELIMINARY; PRT; 119 AA.
AC Q80IR8;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
```

```
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Nonstructural protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JUT-Kan;
RA Takahashi K., Kang J., Ohtsishi S., Hino K., Miyakawa H., Mishiro S.;
RT "Full-genome nucleotide sequence of hepatitis E virus recovered from
RT Japanese patients with acute and/or fulminant hepatitis.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB091394; BAC65251.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2.1.
KW Nonstructural protein.
SQ SEQUENCE 119 AA; 12070 MW; AC80D4D642EDB345 CRC64;

Query Match 20.1%; Score 84; DB 12; Length 119;
Best Local Similarity 37.7%; Pred. No. 0.19;
Matches 23; Conservative 11; Mismatches 21; Indels 6; Gaps 3;

QY 1 GTHSLPRPAAPVPLRMQPGPAHPVLSFLRPSMDLVSFAFYSLLPLAPLSPTSPVISP 57
DB 52 GVTGLILSPS--PSPFIQTPS--PMSFHNPGLEALDSRPVPLAPLVGTSPSAPLPP 108
QY 58 V 58
DB 109 V 109

RESULT 10
Q35615 PRELIMINARY; PRT; 995 AA.
AC Q35615;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FOG.
GN ZPEM1 OR FOG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97373824; PubMed=9230307;
RA Teang A.P., Visvader J.E., Turner C.A., Fujiwara Y., Yu C.,
RA Weiss M.J., Crossley M., Orkin S.H.;
RT "FOG, a multitype zinc finger protein, acts as a cofactor for
RT transcription factor GATA-1 in erythroid and megakaryocytic
RT differentiation.";
RL Cell 90:109-119(1997).
DR EMBL; AF006492; AAC53292.1; -.
DR MGD; MGI:1095400; Zfpml.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPR001005; Myb DNA binding.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 5.
DR SMART; SM00355; Znf_C2H2; 5.
DR PROSITE; PS00037; MYB 1; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 2.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 995 AA; 105983 MW; 29325B2815EC88 CRC64;

Query Match 20.1%; Score 84; DB 11; Length 995;
Best Local Similarity 33.3%; Pred. No. 1.8;
Matches 23; Conservative 10; Mismatches 28; Indels 8; Gaps 2;
```

QY 7 PPAAVPPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSVPIPSVGVGRGP-- 64
DB 765 PPAAVPPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSVPIPSVGVGRGP 820
QY 65 -DPDAHVV 72
DB 821 QSFDPATL 829

RESULT 11

Q8RZY1 PRELIMINARY; PRT; 316 AA.
AC Q8RZY1;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE P0034C09.31 protein.
GN P0034C09.31.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=19947;
RN [1]_TaxID=19947;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
XT clone:P0034C09."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003450; BAB84618.1;
DR Gramene; Q8RZY1;
SQ SEQUENCE 316 AA; 34223 MW; 686ABF29A419C259 CRC64;

Query Match 19.7%; Score 82.5; DB 10; Length 316;
Best Local Similarity 37.9%; Pred. NO. 0.75;
Matches 25; Conservative 8; Mismatches 26; Indels 7; Gaps 3;

QY 4 SLLPRAAVPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSVPIPSVGVGRG 63
DB 54 TLLSPTPVVPM-LLPSTRPVFMQPHFLVPA-----LPPSP-QVPOSSLSSLSA 106
QY 64 PDPDAH 69
DB 107 PGSTRH 112

RESULT 12

Q8V730 PRELIMINARY; PRT; 122 AA.
AC Q8V730;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Phosphoprotein.
OS Swine hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=63421;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=swJ570;
RA Okamoto H.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=swJ570;
RX MEDLINE=21616952; PubMed=11741279;
RA Okamoto H., Takahashi M., Nishizawa T., Fukui K., Muramatsu U.,
RA Yoshikawa A.;
RT "Analysis of the complete genome of indigenous swine hepatitis E virus
isolated in Japan."
RL Biochem. Biophys. Res. Commun. 289:929-936 (2001).

DR EMBL; AB073912; BAB79305.1;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2.1.
SQ SEQUENCE 122 AA; 12397 MW; 6BAA9F02321D26A4 CRC64;

Query Match 19.6%; Score 82; DB 12; Length 122;
Best Local Similarity 37.7%; Pred. NO. 0.31;
Matches 23; Conservative 10; Mismatches 22; Indels 6; Gaps 3;

QY 1 GTHSLPRAAVPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSVPIPS 57
DB 55 GVTGLILSPS--PSFIFIQTPS--PPMSFHNFGLELADSRPAFLPGLGATSPSAPPLPP 111
QY 58 V 58
DB 112 V 112

RESULT 13

Q91115 PRELIMINARY; PRT; 122 AA.
AC Q91115;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE ORF3 protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JRAL;
RA Mishiro S.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=JRAL;
RX MEDLINE=21396683; PubMed=11504536;
RA Takahashi K., Iwata K., Watanabe N., Hatahara T., Ohta Y., Baba K.,
RA Mishiro S.;
RT "Full-genome nucleotide sequence of a hepatitis E virus strain that
may be indigenous to Japan."
RL Virology 287:9-12 (2001).
DR EMBL; AP003430; BAB63940.1;
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003384; HEV_ORF2.
DR Pfam; PF02444; HEV_ORF2.1.
SQ SEQUENCE 122 AA; 12313 MW; D6C56A8E54C587DD CRC64;

Query Match 19.6%; Score 82; DB 12; Length 122;
Best Local Similarity 37.7%; Pred. NO. 0.31;
Matches 23; Conservative 10; Mismatches 22; Indels 6; Gaps 3;

QY 1 GTHSLPRAAVPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSVPIPS 57
DB 55 GVTGLILSPS--PSFIFIQTPS--PPMSFHNFGLELADSRPAFLPGLGATSPSAPPLPP 111
QY 58 V 58
DB 112 V 112

RESULT 14

O70474 PRELIMINARY; PRT; 816 AA.
AC O70474;
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

```

DE Neurocan (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Hippocampus;
RA Zachmann-Brand B., Schaller H.C.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF060879; AAC15766.1; -.
DR HSSP; P00740; 1EDM.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005540; F:hyaluronic acid binding; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000152; Axx hydroxyl_S.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR001881; EGF Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00193; Xlink; 1.
DR PRINTS; P01265; LINKMODULE.
DR ProDom; PD000918; Link; 1.
DR SMART; SM00179; EGF CA; 1.
DR SMART; SM00445; LINK; 1.
DR PROSITE; PS00010; ASX-HYDROXYL; 1.
DR PROSITE; PS00041; C-TYPE LECTIN_2; 1.
DR PROSITE; PS00022; EGF 1; 3.
DR PROSITE; PS01186; EGF 2; 1.
DR PROSITE; PS01187; EGF CA; 1.
DR PROSITE; PS01241; LINK; 1.
KW EGF-like domain.
FT NON_TER 1 816
FT SEQUENCE 816 AA; 85578 MW; 957FS917AD10616E CRC64;

Query Match 19.5%; Score 81.5; DB 11; Length 816;
Best Local Similarity 35.0%; Pred. No. 2.6;
Matches 28; Conservative 10; Mismatches 25; Indels 17; Gaps 5;

OY 4 SLLPFAVPLRMQPG---PAHVLFLR-----PSWDLVSFAFSLPLAPLS--PT 51
Db 342 SSIPSEALSAVLSQASPGSGSPDFEIVAMLRAPKWLPHSTLVNPNVSPISLPASP 401
OY 52 SVP-----ISPVSVGRGDPD 67
Db 402 SVPEEQAVRPVSFG-AEDPE 420

RESULT 15
O82761 PRELIMINARY; PRT; 200 AA.
AC O82761;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DE Putative phytocyanin (Putative blue copper-binding protein).
GN AT2G31050.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Toriumi M., Chan A., Chung M., Goldsmith A., Liu A., Smith A.,
RA Vaysberg M., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,
RA Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,
RA Fujii C.F., Mason I.N., Bowman C.D., Shea T.P., Benito M.-I., Town C.D.,
RA Lin X., Kaul S., Rounsley S.D., Shear T.P., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhagen G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
*Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana";
RT Nature 402:761-768 (1999).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
RA Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005311; AAC63847.1; -.
DR EMBL; AC004669; AAM14981.1; -.
DR PIR; H84715; H84715.
DR HSSP; P00303; 2CBP.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR003245; Plicyanin-like.
DR Pfam; PF02298; Cu bind like; 1.
DR ProDom; PD003122; Plicyanin like; 1.
DR SEQUENCE 200 AA; 21475 MW; E669011C997E349C CRC64;

Query Match 19.4%; Score 81; DB 10; Length 200;
Best Local Similarity 34.3%; Pred. No. 0.67;
Matches 23; Conservative 5; Mismatches 23; Indels 16; Gaps 2;

OY 9 PAAPVPLRMQPGPAHVLFLRPSWDLVSFAFSLPLAPLSPTSPVSVGRGDPDA 68
Db 132 FVAAPV-----PQVRRPSSFSFSS-----QSPLAESPVNHPVQYQMGSPAP 175
OY 69 HVAVNLS 75
Db 176 HSAASNS 182

RESULT 16
O9LMQ1 PRELIMINARY; PRT; 1006 AA.
AC O9LMQ1;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE FTH2.17 protein.
GN FTH2.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Liu S.X., Sakano H., Yu G., Etgu P., Lee J., Lenz C., Pham P.,
RA Toriumi M., Chan A., Chung M., Goldsmith A., Liu A., Smith A.,
RA Vaysberg M., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,
RA Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhagen G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
*Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana";
RT Nature 402:761-768 (1999).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
RA Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005311; AAC63847.1; -.
DR EMBL; AC004669; AAM14981.1; -.
DR PIR; H84715; H84715.
DR HSSP; P00303; 2CBP.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR003245; Plicyanin-like.
DR Pfam; PF02298; Cu bind like; 1.
DR ProDom; PD003122; Plicyanin like; 1.
DR SEQUENCE 200 AA; 21475 MW; E669011C997E349C CRC64;

```

RA Ecker J.R., Federspiel N.A., Theologis A.;
RT "The sequence of BAC F7H2 from Arabidopsis thaliana chromosome 1.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC034256; AAF8153.1; -;
DR FIR: G86292; G86292;
SQ SEQUENCE 1006 AA; 103943 MW; C9F849F9930C238D CRC64;

Query Match
Best Local Similarity 31.4%; Score 81; DB 10; Length 1006;
Matches 22; Conservative 5; Mismatches 21; Indels 2; Gaps 2;

QY 6 LRPAAVPLRMQPGAPHVLSFLRPSWDLVSAFYSLPLAPLSPTSVSPISVSGRPPD 65
DB 336 LPPPATLPPLPLPPPP-----SLFVTCSP-----PPPTIIVNGAPP 373

QY 66 PDARVAVNLS 75
DB 374 PCVTCVQVS 383

RESULT 17
Q9IMY0 PRELIMINARY; PRT; 941 AA.
AC Q9IMY0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE EBNA-3A.
OS Cercopithecine herpesvirus 15.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=104228;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=97048062; PubMed=8892903;
RA Franken M., Devergne O., Rosenzweig M., Annis B., Kieff E., Wang F.;
RT "Comparative analysis identifies conserved tumor necrosis factor
RT receptor-associated factor 3 binding sites in the human and simian
RT Epstein-Barr virus oncogene LMFL.";
RL J. Virol. 70:7819-7826(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=99412410; PubMed=10482645;
RA Rivallier P., Quink C., Wang F.;
RT "Strong selective pressure for evolution of an Epstein-Barr virus
RT LMP2B homologue in the rhesus lymphocryptovirus.";
RL J. Virol. 73:8867-8872(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=20304984; PubMed=10846073;
RA Jiang H., Cho Y.-G., Wang F.;
RT "Structural, functional, and genetic comparisons of Epstein-Barr virus
RT nuclear antigen 3A, 3B, and 3C homologues encoded by the rhesus
RT lymphocryptovirus.";
RL J. Virol. 74:5921-5932(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=20440633; PubMed=10970361;
RA Rao P., Jiang H., Wang F.;
RT "Cloning of the rhesus lymphocryptovirus viral capsid antigen and
RT Epstein-Barr virus-encoded small RNA homologues and use in diagnosis
RT of acute and persistent infections.";
RL J. Clin. Microbiol. 38:3219-3225(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=21602573; PubMed=11739708;
RA Rivallier P., Jiang H., Cho Y.-G., Quink C., Wang F.;
RT "Complete Nucleotide Sequence of the Rhesus Lymphocryptovirus: Genetic

RT Validation for an Epstein-Barr Virus Animal Model.";
RL J. Virol. 76:421-426(2002).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RA Moghaddam A., Koch J., Annis B., Wang F.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RA Moghaddam A., Annis B., Wang F.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RA Rivallier P., Quink C., Wang F.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RA Jiang H., Wang F.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [10]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RA Rao P.V., Jiang H., Wang F.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RA Rivallier P., Jiang H., Cho Y.-G., Quink C., Wang F.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY037858; NAF78881.2;
DR GO: GO:0042025; C:host cell nucleus; IEA.
DR GO: GO:0016032; P:viral life cycle; IEA.
DR InterPro: IPR007706; EBNA-3.
DR Pfam: PF05009; EBNA-3; 1.
SQ SEQUENCE 941 AA; 103089 MW; EA4E3DC9BEC19A92 CRC64;

Query Match 18.8%; Score 78.5; DB 12; Length 941;
Best Local Similarity 31.9%; Pred. No. 6.4; Mismatches 8; Indels 23; Gaps 4;
Matches 22; Conservative 8;

QY 7 PRPAAPVPLRMQPGAPHV-----LSFLRPSWDLVSAFYSLPLAPLSPTSV-PISPV 58
DB 580 PRPAAPLPL-SLALLTGPAPETSSSVQVMSYAAPSM-----AFTPQPVVPI 624
QY 59 SVGRGPPDD 67
DB 625 PIROGPPD 633

RESULT 18
Q8UJMS PRELIMINARY; PRT; 122 AA.
AC Q8UJMS;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE ORF3 protein.
OS Hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=12461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JMY-Haw;
RA Mishiro S.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=JMY-Haw;
RX MEDLINE=21996248; PubMed=12001054;

33

QY 60 VGRGPPDPA 68
DB 69 -----PAPPA 73

RESULT 22
Q9VI97
AC Q9VI97 PRELIMINARY; PRT; 269 AA.
ID Q9VI97
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CG31496 protein.
GN NXF4 OR BCNDNA:AT07692 OR CGI4604 OR CG31501.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
(1)
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter B.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Flosser C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Renington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong P.N., Zhong Z.H., Zhou X., Zhou S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
(2)
RP SEQUENCE FROM N.A.
RA Ceiniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,

RA Pacle J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
(3)
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Ceiniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
(4)
RP SEQUENCE FROM N.A.
RA Adams M.D., Ceiniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
(5)
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003672; AAF54028.2; --
DR FlyBase; FBGN0051501; nxf4.
SQ SEQUENCE 269 AA; 27072 MW; D69FSD742788F331 CRC64;
Query Match 18.5%; Score 77.5; DB 5; Length 269;
Best Local Similarity 30.9%; Pred. No. 2.2;
Matches 21; Conservative 9; Mismatches 23; Indels 15; Gaps 2;
QY 7 PRPAVP-----VPLRMQPGPAHPVLS-----FLRSDVLVSFYSLPLAPLSPT 51
DB 121 PAPAIAPSIAGVPVMPQSPVPFPAVSPAAAPANVAPAPAPVAPVAPVAPVAPVAP 180
QY 52 SVPIPSVPS 59
DB 191 VVPVAPVA 188
RESULT 23
Q9L1C8 PRELIMINARY; PRT; 1212 AA.
AC Q9L1C8
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative nitroreductase.
GN SC01554 OR SCL11.10C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
(1)
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
(2)
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
(3)
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

MoL. Microbiol. 21:77-96(1996).
[4]
RN STRAIN=A3(2) / M145;
RC MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL9139109; CAB76073.1; --
DR HSP; Q05403; LDGV.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006119; P:electron transport; IEA.
DR InterPro; IPR000415; Nitroreductase.
DR InterPro; IPR003200; NN:DBI_PRT.
DR InterPro; IPR008281; NN:DBI_PRT_sub.
DR Pfam; PF02277; DBI_PRT; 1; --
DR Pfam; PF08881; Nitroreductase; 1.
DR ProDom; PD009438; NN:DBI_PRT; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 1212 AA; 124281 MW; E898B903F78D235A CRC64;

Query Match 18.5%; Score 77.5; DB 16; Length 1212;
Best Local Similarity 36.7%; Pred. No. 11;
Matches 22; Conservative 7; Mismatches 28; Indels 3; Gaps 2;

QY 9 PAAVPLRMQPGPAHPVLSFLRPSWDLVSAPYSLPLAPLSPTSPISVSGRGPDDA 68
DB 488 PAAVPAAPVAQHPHPAQVQGLPVDGGVPT--TPHLAPTPEAVLVPPTAAE-PEPVA 544

RESULT 24
Q36612
ID O36612 PRELIMINARY; PRT; 122 AA.
AC O36612;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Swine hepatitis E virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage;
OC Hepatitis E-like viruses.
OX NCBI_TaxID=63421;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Meng;
RX MEDLINE=97420774; PubMed=9275216;
RA Meng X.J., Purcell R.H., Halbur P.G., Lehman J.R., Webb D.M.,
RA Tsareva T.S., Haynes J.S., Thacker B.J., Emerson S.U.;
RT "A novel virus in swine is closely related to the human hepatitis E
RT virus";
RL Proc. Natl. Acad. Sci. U.S.A. 94:9860-9865(1997).
RV [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Meng;
RX MEDLINE=99030877; PubMed=9811705;
RA Meng X.J., Halbur P.G., Shapiro M.S., Govindarajan S., Bruna J.D.,
RA Mushahwar I.K., Purcell R.H., Emerson S.U.;
RT "Genetic and experimental evidence for cross-species infection by
RT swine hepatitis E virus";
RL J. Virol. 72:9714-9721(1998).
DR EMBL; AF082843; AAC97209.1; --
DR GO; GO:0013028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR003384; HEV_ORF2.

DR Pfam; PF02444; HEV_ORF2; 1.
KW Hypothetical protein.
SQ SEQUENCE 122 AA; 12466 MW; 77602F9048E7B12A CRC64;

Query Match 18.4%; Score 77; DB 12; Length 122;
Best Local Similarity 36.1%; Pred. No. 1.1; 23; Indels 6; Gaps 3;
Matches 22; Conservative 10; Mismatches 23; Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE C15H11.5 protein.
OS Caenorhabditis elegans.
OC Rukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC Bardin S.C.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology";
RL Science 282:2012-2018(1998).
CC -|- SIMILARITY: CONTAINS 1 SET DOMAIN.
DR EMBL; Z81035; CAB02737.1; --
DR PIR; T19319; T19319.
DR WormPep; C15H11.5; CB08182.
DR GO; GO:0000785; C:chromatin; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003682; F:chromatin binding; IEA.
DR GO; GO:0006333; P:chromatin assembly/disassembly; IEA.
DR InterPro; IPR000953; Chromo.
DR SMART; SM00317; SET; 1.
DR SMART; SM00298; CHROMO; 1.
DR PROSITE; PS0280; SET; 1.
DR SEQUENCE 503 AA; 58532 MW; E85FA66A9C0F4D2E CRC64;

Query Match 18.4%; Score 77; DB 5; Length 503;
Best Local Similarity 32.7%; Pred. No. 4.8;
Matches 17; Conservative 8; Mismatches 17; Indels 10; Gaps 1;

QY 11 AVPVPLRMQPGPAHPVLSFLRPSWDLVSAPYSLPLAPLSPTSPISVSGR 62
DB 298 SIPLERQEPQ-----SGWYILSAMYSLEPNVPLMTGVTGIRPAEVAR 339

RESULT 26
Q94D41
ID Q94D41 PRELIMINARY; PRT; 690 AA.
AC Q94D41;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE P0712B02.14 protein.

01-JUN-2003 (TReMBLrel. 24, Created)
01-JUN-2003 (TReMBLrel. 24, Last sequence update)
01-OCT-2003 (TReMBLrel. 25, Last annotation update)
FLJ20321 protein.
Homo sapiens (human).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RC
TSSUE=uterus;
MEDLINE=22388257; PubMed=12477932;
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan A., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.D., Scheetz T.E.,
Brownstein M.J., Tsfin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton S., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko I., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzyszinski M.I., Skalske U., Smalusz D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
SEQUENCE FROM N.A.
RC
TSSUE=uterus;
Strausberg R.;
Submitted (May-2003) to the EMBL/GenBank/DBJ databases.
EMBL; BC051883; AAH51883.1;
InterPro; IPR007087; znf_C2H2.
Pfam; PF00096; zf-C2H2_4.
SMART; SM00355; Znf_C2H2; 5.
PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
SEQUENCE 1044 AA; 111730 MW; CF4DC236A70CBAE8 CRC64;
Query Match 18.48; Score 77; DB 4; Length 1044;
Best Local Similarity 34.3%; Pred. No. 10;
Matches 24; Conservative 8; Mismatches 24; Indels 14; Gaps 25
Qy 8 RPAAPVPLRMQGPAPHLVSLFRLPSWDLVSAFVSLPLAPLSPTSVFIPSVGSGPDP 67
Db 250 RPSXYDVRGIQKCPAK-----VPTPTSLAPAPLA--SVPSAPSAPGPGPP 295
Qy 68 AHVAVNLSRY 77
Db 296 ASLSFNTPEY 305
RESULT 29
Q09493 PRELIMINARY; PRT; 1110 AA.
AC Q09493;
DT 01-JAN-1999 (TReMBLrel. 09, Created)
DT 01-JAN-1999 (TReMBLrel. 09, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE C33B4.3 protein.
GN C33B4.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.

RA Coles L.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RW [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RY none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RT Science 292:2012-2018(1998).
CC !- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
DR EMBL; Z48367; CAA88324.1; -.
DR FIR; T19673; T19673.
DR HSSP; P31016; 1BFE.
DR WormPep; C33B4.3; CE01508.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00595; PDZ; 1.
DR Pfam; PF00536; SAM; 1.
DR SMART; SM00248; ANK; 2.
DR SMART; SM00228; PDZ; 1.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50106; PDZ; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1110 AA; 122628 MW; 89A21494B8ED43DB CRC64;

Query Match 18.3%; Score 76.5; DB 5; Length 1110;
Best Local Similarity 28.6%; Pred. No. 12;
Matches 24; Conservative 12; Mismatches 33; Indels 15; Gaps 4;

QY 3 HSLPRPAAPVPLRMQ-----PGPAHPVLSFLRPSWDLVSAFYSLP-LAPLSPTSVFIS 56
DB 744 HPSLRPRASTPPIQQQSSIPPPPPPPPHCEPT--WVHVEFTPTSTSSVPPPPPLP 801

QY 57 PVSVGRGPD-----ARVAVN 73
DB 802 PISGAPPPPPPPPPPGGLMHVAAS 825

RESULT 30

Q9P6R1
ID Q9P6R1 PRELIMINARY; PRT; 309 AA.
AC Q9P6R1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative proline-rich protein verprolin, possibly involved in
DE cytoskeletal organization and cellular growth, actin cytoskeleton-
DE associated.
GN SPBC1387.09.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Wood V., Rajandream M.A., Barrell B.G., Skelton J., Churcher C.M.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL354632; CAB99884.1; -.
DR GenDB_Spombe; SPBC1387.09; -.
DR InterPro; IPR003124; WH2.
DR Pfam; PF02205; WH2; 1.
DR SMART; SM00246; WH2; 1.
SQ SEQUENCE 309 AA; 31224 MW; 47B68086BD93E7D5 CRC64;

Query Match 18.2%; Score 76; DB 3; Length 309;

Best Local Similarity 37.1%; Pred. No. 3.6;
Matches 26; Conservative 6; Mismatches 34; Indels 4; Gaps 2;

QY 7 PRPAAPVPLRMQPGPAHPVLSFLRPSWDLVSAFYSLP-LAPLSPTSVFISVSVGRGPD 66
DB 152 PSPASAP-PI---PSKAPPFSSSLPPPAQPAAPVKSPSPSPSPSPSPSPSPSPSP 207

QY 67 DAHVAVNLSR 76
DB 208 QAPVANTSSR 217

Search completed: July 4, 2004, 04:20:55
Job time : 19.1185 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2004, 04:09:33 ; Search time 86.6606 Seconds
(without alignments)
1366.103 Million cell updates/sec

Title: US-09-506-079H-12
Perfect score: 2287
Sequence: 1 MELAALCRWGLLLALLPFGA.....VGRGPDPAHVAVNLSRYEG 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 90 summaries

Database :
1: A_Geneseq_29Jan04.*
2: Geneseqp1980s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2191	95.4	419	4	Aae09213 Human p68
2	2176	95.1	419	4	Aae09212 Human p68
3	2174	95.1	419	4	Aae09205 Human p68
4	2172	95.0	419	4	Aae09207 Human p68
5	2172	95.0	419	4	Aae09209 Human p68
6	2171	94.9	419	5	Aae20348 Human tru
7	2170	94.9	419	4	Aae09203 Human p68
8	2169	94.8	419	4	Aae09210 Human p68
9	2169	94.8	419	4	Aae09208 Human p68
10	2169	94.8	419	4	Aae09181 Human p68
11	2168	94.8	419	4	Aae09211 Human p68
12	2168	94.8	419	4	Aae09206 Human p68
13	2167	94.8	419	4	Aae09204 Human p68
14	2164	94.6	419	4	Aae09216 Human p68
15	2161	94.5	419	4	Aae09202 Human p68
16	2161	94.5	419	4	Aae09200 Human p68
17	2158	94.4	420	3	Aay97240 Truncated
18	2157	94.3	419	4	Aae09183 Human p68
19	2155	94.2	419	4	Aae09214 Human p68
20	2155	94.2	419	4	Aae09215 Human p68
21	1878	82.1	645	4	AAB60408 Human Erb
22	1878	82.1	645	4	AAB61593 Human Erb
23	1878	82.1	645	5	ABG70753 Human Erb
24	1878	82.1	645	8	Ad871462 Human Erb
25	1878	82.1	653	3	AAB21200 Extracell

ALIGNMENTS

RESULT 1
AAE09213
ID AAE09213 standard; protein; 419 AA.

XX AAE09213;
 XX
 DT 15-NOV-2001 (first entry)
 XX
 DE Human p68HER-2 generic protein variant 11.
 XX
 KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
 KW ECDIIIA; variant.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX Region 1..340
 XX /note= "Identical to N-terminal region of p185HER-2"
 XX Misc-difference 124
 XX /note= "Represented as Agn in the parent sequence shown
 XX in the specification"
 XX Misc-difference 125
 XX /note= "Represented as Agn in the parent sequence shown
 XX in the specification"
 XX Domain 341..419
 XX /label= ECDIIIA variant
 XX /note= "Extracellular domain IIIa variant"
 XX Misc-difference 342
 XX /label= Unknown
 XX Misc-difference 345
 XX /label= Unknown
 XX Misc-difference 346
 XX /note= "p68HER-2 generic sequence (AAE09181) Xaa
 XX substituted with Leu"
 XX Misc-difference 356
 XX /label= Unknown
 XX Misc-difference 358
 XX /label= Unknown
 XX Misc-difference 361
 XX /label= Unknown
 XX Misc-difference 376
 XX /label= Unknown
 XX Misc-difference 394
 XX /label= Unknown
 XX Misc-difference 404
 XX /label= Unknown
 XX Misc-difference 413
 XX /note= "p68HER-2 generic sequence (AAE09181) Xaa
 XX substituted with Asn"
 XX
 XX WO200161356-A1.
 XX
 XX 23-AUG-2001.
 XX
 XX 16-FEB-2001; 2001WO-US005327.
 XX
 XX 16-FEB-2001; 2000US-00506079.
 XX
 XX (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 XX Clinton G, Henner WD, Evans A;
 XX
 XX WPI; 2001-529934/58.
 XX
 XX New polypeptide, which binds to the extracellular domain of HER-2 for the
 XX treatment of hard tumors.
 XX
 XX Example 11; Page; 61pp; English.
 XX
 XX The invention relates to novel HER-2 (herstatin-2) antagonist
 XX particularly a polypeptide that binds to the extracellular domain (ECD)
 XX of HER-2 at a site that is different from the binding site of humanised
 XX antibody, Herceptin, at an affinity of at least 10⁸. The present
 XX invention is based upon the initial discovery of an alternative HER-2
 XX mRNA transcript with 274 bp insert of intron 8. The translation product

CC of the alternative transcript is a truncated HER-2 protein designated
 CC p68HER-2 which lacks the transmembrane and intracellular domains of
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
 CC nucleic acids encoding these are useful to treat, diagnose and identify
 CC solid tumours. The present sequence is human p68HER-2 generic protein
 CC containing ECDIIIA variant sequence. Note: The present sequence is not
 CC shown in the specification but is derived from p68HER-2 generic sequence
 CC (SEQ ID NO:2) shown in the sequence listing (AAE09181)
 XX
 XX Sequence 419 AA;
 XX
 XX Query Match 95.4%; Score 2181; DB 4; Length 419;
 XX Best Local Similarity 95.9%; Pred. NO. 2.1e-165;
 XX Matches 402; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
 QY 1 MELAALCRWGLLLALLPPGAASTQVCTGDMKLRLPASPETHLDMRLHYQGCWVQGNL 60
 DB 1 MELAALCRWGLLLALLPPGAASTQVCTGDMKLRLPASPETHLDMRLHYQGCWVQGNL 60
 QY 61 ELTYLPTNASLSFLQDIQEVQGVLIHQAQVQLQRLRIVRGTLQFEDNYALAVLDNG 120
 DB 61 ELTYLPTNASLSFLQDIQEVQGVLIHQAQVQLQRLRIVRGTLQFEDNYALAVLDNG 120
 QY 121 DPLNNTTPTVGASPGGLRELQRLSLEILKGGVLIQRNPQLCYQDTILWKDIFHKNNQLA 180
 DB 121 DPLNNTTPTVGASPGGLRELQRLSLEILKGGVLIQRNPQLCYQDTILWKDIFHKNNQLA 180
 QY 181 LTLIDTNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCAAGCACKGKPLPTDCCHEQC 240
 DB 181 LTLIDTNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCAAGCACKGKPLPTDCCHEQC 240
 QY 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDTPESMPNPEGRYTFGASCVTACP 300
 DB 241 AAGCTGPKHSDCLACLFHFNHSGICELHCPALVTYNTDTPESMPNPEGRYTFGASCVTACP 300
 QY 301 YNVLSTDVGSCTLVCPHMQEVTAEQTCRCCKPCARGTHSLLPRAAIVPVLWOP 360
 DB 301 YNVLSTDVGSCTLVCPHMQEVTAEQTCRCCKPCARGTHSLLPRAAIVPVLWOP 360
 QY 361 GPAHPVLSFLRPSWDLVSFAFYSPLAPLSPTSPVSPVSGRGPDPDAHVAVNLSRYEG 419
 DB 361 GPAHPVLSFLRPSWDLVSFAFYSPLAPLSPTSPVSPVSGRGPDPDAHVAVNLSRYEG 419
 XX
 XX RESULT 2
 XX AAE09212
 XX ID AAE09212 standard; protein; 419 AA.
 XX AC AAE09212;
 XX DT 15-NOV-2001 (first entry)
 XX DE Human p68HER-2 generic protein variant 10.
 XX
 XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
 XX solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
 XX ECDIIIA; variant.
 XX
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX Region 1..340
 XX /note= "Identical to N-terminal region of p185HER-2"
 XX Misc-difference 124
 XX /note= "Represented as Agn in the parent sequence shown
 XX in the specification"
 XX Misc-difference 125
 XX /note= "Represented as Agn in the parent sequence shown
 XX in the specification"
 XX Domain 341..419
 XX /label= ECDIIIA variant

FT /note= "Extracellular domain IIIa variant"
FT Misc-difference 342 /label= Unknown
FT Misc-difference 345 /label= Unknown
FT Misc-difference 346 /label= Unknown
FT Misc-difference 346 /label= Unknown
FT Misc-difference 356 /label= Unknown
FT Misc-difference 358 /label= Unknown
FT Misc-difference 361 /label= Unknown
FT Misc-difference 376 /label= Unknown
FT Misc-difference 394 /label= Unknown
FT Misc-difference 404 /label= Unknown
FT Misc-difference 413 /label= Unknown
FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Asn"
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UVOR-) UNIV OREGON HEALTH SCI.
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX Example 11; Page: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 10⁸. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDIIIa.
XX The ECDIIIa-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIa variant sequence. Note: The present sequence is not
XX shown in the specification but is derived from p68HER-2 generic sequence
XX (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
SQ Sequence 419 AA;

Query Match 95.1%; Score 2176; DB 4; Length 419;
Best Local Similarity 95.7%; Pred. No. 5.3e-165;
Matches 401; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLALLPFGAASQVCTGDKLRLPASPEHLDMLRHLYQGQVQGNL 60
Db 1 MELAALCRWGLLALLPFGAASQVCTGDKLRLPASPEHLDMLRHLYQGQVQGNL 60
QY 61 ELTYLPTNLSFLQDIQEVQGVYLIAHNVQVQVPLQRLIRVGTQLFEDNYALVLDNG 120
Db 61 ELTYLPTNLSFLQDIQEVQGVYLCAHNVQVQVPLQRLIRVGTQLFEDNYALVLDNG 120

QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
Db 121 DFLXXTTPVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPCKSGRCWGESSEDCQSLTRITVCAGGCARCKGPLPTDCCHEQC 240
Db 181 LTLIDTNRSRACHPCSPCKSGRCWGESSEDCQSLTRITVCAGGCARCKGPLPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMNPETGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMNPETGRYTFGASCVTACP 300
QY 301 YNVLTSDVGSCTLVCPHNOEVTAEQTCRCEKSKPCARGTHSLPRPAAPVPLRMQP 360
Db 301 YNKLSTDVGSCTLVCPHNOEVTAEQTCRCEKSKPCARGTHSLPRPAAPVPLRMQP 360
QY 361 GPAHPVLSPLRPSWDLWSAFYSLPLAPLSPVTPSPVSGRGPDPDAHVANLSRYEG 419
Db 361 XPAHPVLSPLRPSWDLWSAFYSLPLAPLSPVTPSPVSGRGPDPDAHVANLSRYEG 419
RESULT 3
AAE09205
ID AAE09205 standard; protein; 419 AA.
XX
XX AAE09205;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 generic protein variant 3.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
XX ECDIIIa; variant.
XX
XX Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FH Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= ECDIIIa variant
FT /note= "Extracellular domain IIIa variant"
FT Misc-difference 342
FT /label= Unknown
FT Misc-difference 345
FT /label= Unknown
FT Misc-difference 346
FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Leu"
FT Misc-difference 356
FT /label= Unknown
FT Misc-difference 358
FT /label= Unknown
FT Misc-difference 361
FT /label= Unknown
FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown
XX WO200161356-A1.
XX
XX

PD 23-AUG-2001.
XX
PF 16-FEB-2001; 2001WO-US005327.
XX
PR 16-FEB-2000; 2000US-00506079.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Clinton G, Henner WD, Evans A;
XX
DR WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
PT treatment of hard tumors.
XX
XX Example 11; Page; 61pp; English.
XX
CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10⁻⁸. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIa.
CC The ECDIIa-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIa variant sequence. Note: The present sequence is not
CC shown in the specification but is derived from p68HER-2 generic sequence
CC (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
SQ Sequence 419 AA;
Query Match 95.1%; Score 2174; DB 4; Length 419;
Best Local Similarity 95.7%; Pred. No. 7, 7e-165;
Matches 401; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 1 MELAALCRWGLLLALLPFGAASQTCTGDKLRLPASPETHLDMRLHYGCGVQGNL 60
DB 1 MELAALCRWGLLLALLPFGAASQTCTGDKLRLPASPETHLDMRLHYGCGVQGNL 60
QY 61 ELTYLPTNASLFLQDIOEVQGYVLCANQVRQVQLRIVRGTLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLFLQDIOEVQGYVLCANQVRQVQLRIVRGTLFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTGASPGGLRELOLRSLTEILKGGVLIQRNPQLCYQDTILWKDIFHKNQLA 180
DB 121 DPLNNTTPTVTGASPGGLRELOLRSLTEILKGGVLIQRNPQLCYQDTILWKDIFHKNQLA 180
QY 181 LTLIDTNRSRACHPCSPCKGSRGSESDCSQSLTRTVCGGCKARCKGPLPTDCCHQOC 240
DB 181 LTLIDTNRSRACHPCSPCKGSRGSESDCSQSLTRTVCGGCKARCKGPLPTDCCHQOC 240
QY 241 AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNLYSTDVGSCTLVCPHINQVEATDGTORCEKSKPCARGTHSLPRPAVPLKQCP 360
DB 301 YNLYSTDVGSCTLVCPHINQVEATDGTORCEKSKPCARGTHSLPRPAVPLKQCP 360
QY 361 GPAHVLSPRLPSWDLVSFAFSLPLAFSLPTSPVSPVSGRGDPDPAHVAVNLSRYEG 419
DB 361 XPAHVLSPRLPSWDLVSFAFSLPLAFSLPTSPVSPVSGRGDPDPAHVAVNLSRYEG 419
RESULT 4
AAE09207
ID AAE09207 standard; protein; 419 AA.
XX

AAE09207;
XX
DT 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 generic protein variant 5.
XX
KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIa; variant.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= "ECDIIa variant
FT /note= "Extracellular domain IIIa variant"
FT Misc-difference 342
FT /label= Unknown
FT Misc-difference 345
FT /label= Unknown
FT Misc-difference 346
FT /label= Unknown
FT Misc-difference 356
FT /label= Unknown
FT Misc-difference 358
FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Leu"
FT Misc-difference 361
FT /label= Unknown
FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown
XX
XX WO200161356-A1.
XX
PD 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
FT treatment of hard tumors.
XX
XX Example 11; Page; 61pp; English.
XX
CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10⁻⁸. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIa.
CC The ECDIIa-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIa variant sequence. Note: The present sequence is not
CC shown in the specification but is derived from p68HER-2 generic sequence
CC (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX

DATE: 11/20/2009

ט
כ
ז
ח

Db 121 DPLXKTTTPTVTGASPGGLRELQLRSLTECLKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
Qy 181 LTLDITNRSRACHPCSPCKGSRGWGSESDCQSLTRTVACGGCARCKGGLPTDCCHEQC 240
Db 181 LTLDITNRSRACHPCSPCKGSRGWGSESDCQSLTRTVACGGCARCKGGLPTDCCHEQC 240
Qy 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSKPCARGTHSLLPDPAAPVPLRMQP 360
Db 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSKPCARGTHSLLPDPAAPVPLRMQP 360
Qy 361 GPAHPVLSFLRPSNDLYSAFYSLPLAPLSPTSPVSPVSGRGPDPDAHVAVNLSRYEG 419
Db 361 XPAHPVLSFLRPSNDLYSAFYSLPLAPLSPTSPVSPVSGRGPDPDAHVAVNLSRYEG 419

RESULT 6
AAE20348
ID AAE20348 standard; protein; 419 AA.
XX AC AAE20348;
XX DT 18-JUN-2002 (first entry)
XX DE Human truncated HER2 protein.
XX KW Human; tumour; endothelial growth factor receptor; EGFR; cytostatic;
KW herstatin; HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung;
XX KW colon; glial cell tumour; cell growth.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
FT Misc-difference 342 /label= Thr, Ser
FT Misc-difference 345 /label= Leu, Pro
FT Misc-difference 346 /label= Pro, Leu
FT Misc-difference 356 /label= Leu, Gln
FT Misc-difference 358 /label= Met, Leu
FT Misc-difference 361 /label= Gly, Asp, Ala, Val
FT Misc-difference 376 /label= Leu, Ile
FT Misc-difference 394 /label= Pro, Arg
FT Misc-difference 404 /label= Pro, Leu
FT Misc-difference 413 /label= Asp, Asn
XX WO200214470-A2.
XX PD 21-FEB-2002.
XX PP 14-AUG-2001; 2001WO-US025502.
XX PR 14-AUG-2000; 2000US-00638834.
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX PI Clinton GM;
XX DR WPI; 2002-269185/31.
XX PT Treating solid tumor characterized by expression of endothelial growth
PT factor receptor, involves administering recombinant herstatin that binds

PT to extracellular domain of the endothelial growth factor receptor.
XX Claim 1; Page 78-80; 82pp; English.
XX The present invention relates to a method for treating a solid tumour
CC characterised by endothelial growth factor receptor (EGFR) expression.
CC The method involves administering an agent that binds to an extracellular
CC domain (ECD) of EGFR. The invention also relates to a naturally occurring
CC inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-
CC expression of herstatin with p185HER2 causes a striking reduction in cell
CC growth that corresponds with suppression of p185 autophosphorylation. The
CC method or a pharmaceutical composition is useful for treating a solid
CC tumour (selected from squamous cell carcinoma, lung carcinoma, colon
CC carcinoma and glial cell tumour) characterised by EGFR expression. The
CC present sequence is human truncated HER2 protein that lacks transmembrane
CC and intracellular domains
XX SQ Sequence 419 AA;
Query Match 94.9%; Score 2171; DB 5; Length 419;
Best Local Similarity 95.5%; Pred. No. 1.3e-164;
Matches 400; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Qy 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPETHLDMLRLHYQGQVQGNL 60
Db 1 MELAALCRWGLLLALLPPGAASQVCTGTDCKLRPASPETHLDMLRLHYQGQVQGNL 60
Qy 61 ELTYLPTNASLSFLQDIQEVQGYVLIAHQVQVPLQRLRIVRGTLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIQEVQGYVLCAHQVQVPLQRLRIVRGTLFEDNYALAVLDNG 120
Qy 121 DPLMNTPTVTGASPGGLRELQLRSLTECLKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
Db 121 DPLRRTPTVTGASPGGLRELQLRSLTECLKGGVLIQRNPOLCYQDTILWKDIFHKNNOLA 180
Qy 181 LTLDITNRSRACHPCSPCKGSRGWGSESDCQSLTRTVACGGCARCKGGLPTDCCHEQC 240
Db 181 LTLDITNRSRACHPCSPCKGSRGWGSESDCQSLTRTVACGGCARCKGGLPTDCCHEQC 240
Qy 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSKPCARGTHSLLPDPAAPVPLRMQP 360
Db 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSKPCARGTHSLLPDPAAPVPLRMQP 360
Qy 361 GPAHPVLSFLRPSNDLYSAFYSLPLAPLSPTSPVSPVSGRGPDPDAHVAVNLSRYEG 419
Db 361 XPAHPVLSFLRPSNDLYSAFYSLPLAPLSPTSPVSPVSGRGPDPDAHVAVNLSRYEG 419

RESULT 7
AAE09203
ID AAE09203 standard; protein; 419 AA.
XX AC AAE09203;
XX DT 15-NOV-2001 (first entry)
XX DE Human p68HER-2 generic protein variant 1.
XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
FT Region 1..340
FT Misc-difference 124 /note= "Identical to N-terminal region of p185HER-2"
FT /note= "Represented as Agn in the parent sequence shown

FT in the specification"
FT Misc-difference 125 /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= ECDIIIIa variant
FT /note= "Extracellular domain IIIa variant"
FT Misc-difference 342 /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Ser"
FT Misc-difference 345 /label= Unknown
FT Misc-difference 346 /label= Unknown
FT Misc-difference 356 /label= Unknown
FT Misc-difference 358 /label= Unknown
FT Misc-difference 361 /label= Unknown
FT Misc-difference 376 /label= Unknown
FT Misc-difference 394 /label= Unknown
FT Misc-difference 404 /label= Unknown
FT Misc-difference 413 /label= Unknown
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX Example 11; Page; 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 10⁸. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p85HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIIa.
XX The ECDIIIIa-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIIa variant sequence. Note: The present sequence is not
XX shown in the specification but is derived from p68HER-2 generic sequence
XX (SEQ ID NO:2) shown in the sequence listing (AAE09181).
XX
XX Sequence 419 AA;
XX
XX Query Match 94.9%; Score 2170; DB 4; Length 419;
XX Best Local Similarity 95.5%; Pred. No. 1.6e-164;
XX Matches 400; Conservative 1; Mismatches 18; Indels 0; Gaps 0;
XX
XX 1 MELAALCRWGLLLALLPFGAASTQVCTGTDCKLRLPASPETHLDMRLHLYGQCQVVGQNL 60
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 1 MELAALCRWGLLLALLPFGAASTQVCTGTDCKLRLPASPETHLDMRLHLYGQCQVVGQNL 60
Qy 61 ELTYLPTNASLSFLQDIQEVQGYVLIHAHQVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIQEVQGYVLIHAHQVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
Qy 121 DPLANNTPVTGASPGGLRRLRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
Db 121 DPLAXTTPTVTGASPGGLRRLRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
Qy 181 LTLIDTNRSRACHPCSPCKSGRCWGESSEDCQSLTRTVGAGGCARCKGPLTDCCHQC 240
Db 181 LTLIDTNRSRACHPCSPCKSGRCWGESSEDCQSLTRTVGAGGCARCKGPLTDCCHQC 240
Qy 241 AAGCTGPKXSDCLACLFHNSGICELHCPALVTYNTDTFESKPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKXSDCLACLFHNSGICELHCPALVTYNTDTFESKPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTGVSGCTLVCPHLNQEVTRAEDGTQRCCKSKPCARGTHSLPRPAAVPLRMOP 360
Db 301 YNKLSTGVSGCTLVCPHLNQEVTRAEDGTQRCCKSKPCARGSHSXXPPPAVFPVPRXQP 360
Qy 361 GPAHPVLSFLRPSMDLVSAFYSLLPLAPLSPTSPISVSVGRGPPDPAHVAVNLGRYEG 419
Db 361 XPAHPVLSFLRPSMDLVSAFYSLLPLAPLDPSTSVXISPSVSVGRGKDPDAFAVAVXLSRYEG 419
RESULT 8
AAE09210
ID AAE09210 standard; protein; 419 AA.
XX
XX AAE09210;
XX
XX 15-NOV-2001 (first entry)
XX Human p68HER-2 generic protein variant 8.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
XX ECDIIIIa; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 1..340
XX /note= "Identical to N-terminal region of p185HER-2"
XX Misc-difference 124 /note= "Represented as Agn in the parent sequence shown
XX in the specification"
XX Misc-difference 125 /note= "Represented as Agn in the parent sequence shown
XX in the specification"
XX Domain 341..419 /label= ECDIIIIa variant
XX /note= "Extracellular domain IIIa variant"
XX Misc-difference 342 /label= Unknown
XX Misc-difference 345 /label= Unknown
XX Misc-difference 346 /label= Unknown
XX Misc-difference 356 /label= Unknown
XX Misc-difference 358 /label= Unknown
XX Misc-difference 361 /label= Unknown
XX Misc-difference 376 /label= Unknown
XX Misc-difference 394 /label= Unknown
XX /note= "p68HER-2 generic sequence (AAE09181) Xaa
XX substituted with Arg"
XX Misc-difference 404

FT /label= Unknown
FT Misc-difference 413
XX /label= Unknown
XX WO200161356-A1.
XX
PD 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX Example 11; Page: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 10⁸. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumors. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence. Note: The present sequence is not
XX shown in the specification but is derived from p68HER-2 generic sequence
XX (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
XX Sequence 419 AA;
XX
XX Query Match 94.8%; Score 2169; DB 4; Length 419;
XX Best Local Similarity 95.5%; Pred. No. 1.9e-164;
XX Matches 400; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
XX
XX 1 MELAALCRWGLLALLPPGAASVCTGTDKMLRLPASPEYHLMRLHYGQCVVQGNL 60
XX
XX 1 MELAALCRWGLLALLPPGAASVCTGTDKMLRLPASPEYHLMRLHYGQCVVQGNL 60
XX
XX 61 ELTYLPTNASLFLQDIQEVQGYVLIANNQVRQVPLQRLIRVGRQLFEDNYALAVLNG 120
XX
XX 61 ELTYLPTNASLFLQDIQEVQGYVLIANNQVRQVPLQRLIRVGRQLFEDNYALAVLNG 120
XX
XX 121 DPLNNTPTVGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
XX
XX 121 DPLNNTPTVGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
XX
XX 181 LTLIDNRSRACHPCSPCKGSRGWGESSEDCQSLTRIVTCAGGCARCKGPLPTDCCHQC 240
XX
XX 181 LTLIDNRSRACHPCSPCKGSRGWGESSEDCQSLTRIVTCAGGCARCKGPLPTDCCHQC 240
XX
XX 241 AAGCTGPKHSDCLAGLHNSGICELHCPALVTYNTDTFESCPNPEGRYTTGASCVTACP 300
XX
XX 241 AAGCTGPKHSDCLAGLHNSGICELHCPALVTYNTDTFESCPNPEGRYTTGASCVTACP 300
XX
XX 301 YNKLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARGTHSLPRPAAPVPLRMQP 360
XX
XX 301 YNKLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARGTHSLPRPAAPVPLRMQP 360
XX
XX 361 GPAHPVLSRLSWDLVSFYLPLAPLPTSPVTPSPVSVGKGGPPDPAHVAUNLSRYEG 419
XX
XX 361 XPAHPVLSRLSWDXVSFYLPLAPLPTSPVTPSPVSVGKGGPPDPAHVAUNLSRYEG 419

RESULT 9
AAE09208
ID AAE09208 standard; protein; 419 AA.
XX
XX AAE09208;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 generic protein variant 6.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
XX ECDIIIA; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 1..340
XX /note= "Identical to N-terminal region of p185HER-2"
XX
XX Misc-difference 124
XX /note= "Represented as Agn in the parent sequence shown
XX in the specification"
XX
XX Misc-difference 125
XX /note= "Represented as Agn in the parent sequence shown
XX in the specification"
XX
XX Domain 341..419
XX /label= ECDIIIA variant
XX /note= "Extracellular domain IIIa variant"
XX
XX Misc-difference 342
XX /label= Unknown
XX
XX Misc-difference 345
XX /label= Unknown
XX
XX Misc-difference 346
XX /label= Unknown
XX
XX Misc-difference 356
XX /label= Unknown
XX
XX Misc-difference 358
XX /label= Unknown
XX
XX Misc-difference 361
XX /label= Asp, Ala, Val
XX
XX Misc-difference 376
XX /label= Unknown
XX
XX Misc-difference 394
XX /label= Unknown
XX
XX Misc-difference 404
XX /label= Unknown
XX
XX Misc-difference 413
XX /label= Unknown
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX Example 11; Page: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 10⁸. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumors. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence. Note: The present sequence is not
XX shown in the specification but is derived from p68HER-2 generic sequence
XX (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
XX Sequence 419 AA;
XX
XX Query Match 94.8%; Score 2169; DB 4; Length 419;
XX Best Local Similarity 95.5%; Pred. No. 1.9e-164;
XX Matches 400; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
XX
XX 1 MELAALCRWGLLALLPPGAASVCTGTDKMLRLPASPEYHLMRLHYGQCVVQGNL 60
XX
XX 1 MELAALCRWGLLALLPPGAASVCTGTDKMLRLPASPEYHLMRLHYGQCVVQGNL 60
XX
XX 61 ELTYLPTNASLFLQDIQEVQGYVLIANNQVRQVPLQRLIRVGRQLFEDNYALAVLNG 120
XX
XX 61 ELTYLPTNASLFLQDIQEVQGYVLIANNQVRQVPLQRLIRVGRQLFEDNYALAVLNG 120
XX
XX 121 DPLNNTPTVGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
XX
XX 121 DPLNNTPTVGASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
XX
XX 181 LTLIDNRSRACHPCSPCKGSRGWGESSEDCQSLTRIVTCAGGCARCKGPLPTDCCHQC 240
XX
XX 181 LTLIDNRSRACHPCSPCKGSRGWGESSEDCQSLTRIVTCAGGCARCKGPLPTDCCHQC 240
XX
XX 241 AAGCTGPKHSDCLAGLHNSGICELHCPALVTYNTDTFESCPNPEGRYTTGASCVTACP 300
XX
XX 241 AAGCTGPKHSDCLAGLHNSGICELHCPALVTYNTDTFESCPNPEGRYTTGASCVTACP 300
XX
XX 301 YNKLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARGTHSLPRPAAPVPLRMQP 360
XX
XX 301 YNKLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARGTHSLPRPAAPVPLRMQP 360
XX
XX 361 GPAHPVLSRLSWDLVSFYLPLAPLPTSPVTPSPVSVGKGGPPDPAHVAUNLSRYEG 419
XX
XX 361 XPAHPVLSRLSWDXVSFYLPLAPLPTSPVTPSPVSVGKGGPPDPAHVAUNLSRYEG 419

Query Match 94.8%; Score 2169; DB 4; Length 419;
Best Local Similarity 95.5%; Pred. No. 1.9e-164;
Matches 400; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MELAALCRWGLLLALLPPGAASSTQVCTGDMKRLPASPETHDMLRHLVQGCQVQGNL 60
Db 1 MELAALCRWGLLLALLPPGAASSTQVCTGDMKRLPASPETHDMLRHLVQGCQVQGNL 60

Qy 61 ELTYLPTNASLSFLQDIQEVQGYVLIHQNVRQVPLQRLRIVRGTLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIQEVQGYVLIHQNVRQVPLQRLRIVRGTLFEDNYALAVLDNG 120

Qy 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHNKNOLA 180
Db 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHNKNOLA 180

Qy 181 LTLIDTNRSRACHPCSPCKGRCWGESSEDCQSLTRTVCAAGCARCKGELPTDCCHQC 240
Db 181 LTLIDTNRSRACHPCSPCKGRCWGESSEDCQSLTRTVCAAGCARCKGELPTDCCHQC 240

Qy 241 AAGCTGPKGSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKGSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

Qy 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARGTHSLPRPAAVPLRMQP 360
Db 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARGTHSLPRPAAVPLRMQP 360

Qy 361 GPAHPVLSFLRPSWDXSAFVSLPLADPTSVISVSVSGRGDPAHVAVNLRYEG 419
Db 361 XPAHPVLSFLRPSWDXSAFVSLPLADPTSVISVSVSGRGDPAHVAVNLRYEG 419

RESULT 11
AAE09211
ID AAE09211 standard; protein; 419 AA.
AC AAE09211;
XX
DT 15-NOV-2001 (first entry)
XX
DB Human p68HER-2 generic protein variant 9.
XX
KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
OS Homo sapiens.
FH
FH Key Location/Qualifiers
FH Region 1..340
FT
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= "ECDIIIA variant
FT /note= "Extracellular domain IIIA variant"
FT Misc-difference 342
FT /label= Unknown
FT Misc-difference 345
FT /label= Unknown
FT Misc-difference 346
FT /label= Unknown
FT Misc-difference 356
FT /label= Unknown
FT Misc-difference 358
FT /label= Unknown
FT Misc-difference 361

FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Leu"
FT Misc-difference 413
FT /label= Unknown
XX
PN WO200161356-A1.
XX
PD 23-AUG-2001.
XX
PF 16-FEB-2001; 2001NO-US005327.
XX
PR 16-FEB-2000; 2000US-00506079.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Clinton G, Henner WD, Evans A;
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
PS Example 11; Page: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 10⁸. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence. Note: The present sequence is not
XX shown in the specification but is derived from p68HER-2 generic sequence
XX (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
XX Sequence 419 AA;
XX
Query Match 94.8%; Score 2169; DB 4; Length 419;
Best Local Similarity 95.5%; Pred. No. 2.3e-164;
Matches 400; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MELAALCRWGLLLALLPPGAASSTQVCTGDMKRLPASPETHDMLRHLVQGCQVQGNL 60
Db 1 MELAALCRWGLLLALLPPGAASSTQVCTGDMKRLPASPETHDMLRHLVQGCQVQGNL 60

Qy 61 ELTYLPTNASLSFLQDIQEVQGYVLIHQNVRQVPLQRLRIVRGTLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIQEVQGYVLIHQNVRQVPLQRLRIVRGTLFEDNYALAVLDNG 120

Qy 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHNKNOLA 180
Db 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHNKNOLA 180

Qy 181 LTLIDTNRSRACHPCSPCKGRCWGESSEDCQSLTRTVCAAGCARCKGELPTDCCHQC 240
Db 181 LTLIDTNRSRACHPCSPCKGRCWGESSEDCQSLTRTVCAAGCARCKGELPTDCCHQC 240

Qy 241 AAGCTGPKGSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKGSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

QY 301 YNYLTDVGSCTLVCPHNEQVTAEDGTQRCCKSPKARGTHSLPRPAAVPPVLRMQP 360
DB 301 YNYLTDVGSCTLVCPHNEQVTAEDGTQRCCKSPKARGTHSLPRPAAVPPVLRMQP 360
QY 361 GRAHPVLSFLRPSWDVSAFYSLPLAPLSTSPISVSVGRGDDPAHVAVNLSRYEG 419
DB 361 XPAHPVLSFLRPSWDVSAFYSLPLAPLSTSPISVSVGRGDDPAHVAVNLSRYEG 419

RESULT 12
AAE09206
ID AAE09206 standard; protein; 419 AA.
XX AAE09206;
XX
DT 15-NOV-2001 (first entry)
DE Human p68HER-2 generic protein variant 4.
DE
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= ECDIIIA variant
FT /note= "Extracellular domain IIIa variant"
FT Misc-difference 342
FT /label= Unknown
FT Misc-difference 345
FT /label= Unknown
FT Misc-difference 346
FT /label= Unknown
FT Misc-difference 356
FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT substituted with Gln"
FT Misc-difference 358
FT /label= Unknown
FT Misc-difference 361
FT /label= Unknown
FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX

PT New polypeptide, which binds to the extracellular domain of HER-2 for the
PT treatment of hard tumors.
XX
PS Example 11; Page; 61pp; English.
XX
CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10⁸. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise,
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIIA variant sequence. Note: The present sequence is not
CC shown in the specification but is derived from p68HER-2 generic sequence
CC (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX
SQ Sequence 419 AA;
Query Match 94.8%; Score 2168; DB 4; Length 419;
Best Local Similarity 95.5%; Pred. No. 2.3e-164;
Matches 400; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYGQCVVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTGCTCKLRLPASPETHLDMRLHLYGQCVVQGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNOVQVPLQRLRVIRGTQLPFDNYALVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNOVQVPLQRLRVIRGTQLPFDNYALVLDNG 120
QY 121 DPLNNTTQVTCASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTTQVTCASPGGLRELQRLSLTEILKGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
QY 181 LTLIDTNRSRACHPCSPCKGSRGWSSESQSLTQVTCAGGCAKCKGPTDCCHEQC 240
DB 181 LTLIDTNRSRACHPCSPCKGSRGWSSESQSLTQVTCAGGCAKCKGPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLTDVGSCTLVCPHNEQVTAEDGTQRCCKSPKARGTHSLPRPAAVPPVLRMQP 360
DB 301 YNYLTDVGSCTLVCPHNEQVTAEDGTQRCCKSPKARGTHSLPRPAAVPPVLRMQP 360
QY 361 GRAHPVLSFLRPSWDVSAFYSLPLAPLSTSPISVSVGRGDDPAHVAVNLSRYEG 419
DB 361 XPAHPVLSFLRPSWDVSAFYSLPLAPLSTSPISVSVGRGDDPAHVAVNLSRYEG 419

RESULT 13
AAE09204
ID AAE09204 standard; protein; 419 AA.
XX
XX AAE09204;
XX
DT 15-NOV-2001 (first entry)
DE Human p68HER-2 generic protein variant 2.
DE
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA; variant.
XX
XX Homo sapiens.
XX

FH Key Location/Qualifiers
FT Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown in the specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the parent sequence shown in the specification"
FT Domain 341..419
FT /label= ECDIIIA variant
FT /note= "Extracellular Domain IIIa variant"
FT Misc-difference 342
FT /label= Unknown
FT Misc-difference 345
FT /note= "p68HER-2 generic sequence (AAE09181) Xaa substituted with Pro"
FT Misc-difference 346
FT /label= Unknown
FT Misc-difference 356
FT /label= Unknown
FT Misc-difference 358
FT /label= Unknown
FT Misc-difference 361
FT /label= Unknown
FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown
XX WO200161356-A1.
XX 23-AUG-2001.
XX 16-FEB-2001; 2001WO-US005327.
XX 16-FEB-2000; 2000US-00506079.
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX Clinton G, Henner WD, Evans A;
XX WPI; 2001-529934/58.
XX New polypeptide, which binds to the extracellular domain of HER-2 for the treatment of hard tumors.
XX Example 11; Page; 61pp; English.
XX The invention relates to novel HER-2 (herstatin-2) antagonist particularly a polypeptide that binds to the extracellular domain (ECD) of HER-2 at a site that is different from the binding site of humanised antibody, Herceptin, at an affinity of at least 10⁸. The present invention is based upon the initial discovery of an alternative HER-2 mRNA transcript with 274 bp insert of intron 8. The translation product of the alternative transcript is a truncated HER-2 protein designated p68HER-2 which lacks the transmembrane and intracellular domains of p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA. The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the nucleic acids encoding these are useful to treat, diagnose and identify solid tumours. The present sequence is human p68HER-2 generic protein containing ECDIIIA variant sequence. Note: The present sequence is not shown in the specification but is derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX Sequence 419 AA;
XX Query Match 94.8%; Score 2167; DB 4; Length 419;

Best Local Similarity 95.5%; Pred. No. 2.8e-164;
Matches 400; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Qy 1 METAALCRWGLLIALLPPGGAASQTGCTDMKRLPASPETHLDMLRHLHYQGCVVQGNL 60
Db 1 METAALCRWGLLIALLPPGGAASQTGCTDMKRLPASPETHLDMLRHLHYQGCVVQGNL 60
Qy 61 ELTYLPTNASLSFLQDIQEVQGVYLIHAHQVQVPLQRLRIVRGTQLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIQEVQGVYLIHAHQVQVPLQRLRIVRGTQLFEDNYALAVLDNG 120
Qy 121 DPLNNTPTVTGASPGGLREQLRLSLTEILKGGVLIQRPOLCYQDTILMKDIFHKNQLA 180
Db 121 DPLXXTPTVTGASPGGLREQLRLSLTEILKGGVLIQRPOLCYQDTILMKDIFHKNQLA 180
Qy 181 LTLIDNRSRACHPCSPCKGRCWGESSEDCOSLRTVTCAGGCACCKGFLPTDCCHEQC 240
Db 181 LTLIDNRSRACHPCSPCKGRCWGESSEDCOSLRTVTCAGGCACCKGFLPTDCCHEQC 240
Qy 241 AAGCTGPKHSDCLALHFNHSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLALHFNHSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHNOEVTAEQGTORCEKSCPCARGTHSLPRPAAPVPLRMQP 360
Db 301 YNKLSTDVGSCTLVCPHNOEVTAEQGTORCEKSCPCARGTHSLPRPAAPVPLRMQP 360
Qy 361 GPAHPVLSFLRPSWDLVSAFYSLLPLAPLSPTSPVSPVSGRGFDDAHVAVXLSRYEG 419
Db 361 XPAHPVLSFLRPSWDLVSAFYSLLPLAPLSPTSPVSPVSGRGFDDAHVAVXLSRYEG 419
RESULT 14
AAE09216
ID AAE09216 standard; protein; 419 AA.
XX AC AAE09216;
XX DT 15-NOV-2001 (first entry)
XX DB Human p68HER-2 generic protein variant 15.
XX KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin; solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2; ECDIIIA; variant.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Region 1..340
XX FT /note= "Identical to N-terminal region of p185HER-2"
XX FT Misc-difference 124
XX /note= "Represented as Agn in the sequence shown in the specification"
XX FT Misc-difference 125
XX /note= "Represented as Agn in the sequence shown in the specification"
XX FT Domain 341..419
XX /label= ECDIIIA variant
XX /note= "Extracellular domain IIIa"
XX FT Misc-difference 342
XX /label= Unknown
XX FT Misc-difference 345
XX /label= Unknown
XX FT Misc-difference 346
XX /label= Unknown
XX FT Misc-difference 356
XX /label= Unknown
XX FT Misc-difference 357
XX /label= Unknown
XX FT Misc-difference 358
XX /label= Unknown
XX FT Misc-difference 361

FT /label= Unknown
FT Misc-difference 371
FT /label= Unknown
FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /note= "p68HER-2 generic sequence (AAE09183) Xaa
FT substituted with Asn"
FT
PN WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX Clinton G, Henner WD, Evans A;
XX PI
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX Example 12; Page; 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 10⁸. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence. Note: The present sequence is not
XX shown in the specification but is derived from p68HER-2 generic sequence
XX (SEQ ID NO:13) shown in the sequence listing (AAE09183)
XX
SQ Sequence 419 AA;

Query Match 94.6%; Score 2164; DB 4; Length 419;
Best Local Similarity 95.2%; Pred. No. 4.8e-164;
Matches 399; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLALLPPGAASQVCTGTDMLRLPASPEHLDMLRHLYQGQVQGNL 60
DB 1 MELAALCRWGLLALLPPGAASQVCTGTDCKRLRLPASPEHLDMLRHLYQGQVQGNL 60
QY 61 ELTYLPTNASLSFLQIQEYQGYVLIHAHQVQVPLQRLRIVRGTLQFDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQIQEYQGYVLIHAHQVQVPLQRLRIVRGTLQFDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQSLTEILKGGVLIQORNPOLCYQDDTLMKDIPHKKNOLA 180
DB 121 DPLXXTPTVTGASPGGLRELQSLTEILKGGVLIQORNPOLCYQDDTLMKDIPHKKNOLA 180
QY 181 LTLIDNRSRACHPCSPCKGSRGWSGESSDCQSLRTVTCAGGCARCKGPELPTDCCHEQC 240
DB 181 LTLIDNRSRACHPCSPCKGSRGWSGESSDCQSLRTVTCAGGCARCKGPELPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHFNHSGICELECPALVYNTDTFESCPNPEGRYTFGASCVTACP 300

Db 241 AAGCTGPKHSDCLACLFHFNHSGICELECPALVYNTDTFESCPNPEGRYTFGASCVTACP 300
QY 301 YYILSTDVGSCTLYVCPLEHNOEVTAEDGTQRCCKSCKPCARGTHSLLPRDAAVPLRMOP 360
Db 301 YNKLTSDVGSCTLYVCPLEHNOEVTAEDGTQRCCKSCKPCARGTHSXXPRPAAVPPVXXXQF 360
QY 361 GPAHPVLSFLRDSMDLVSAFSLPLAPLSPTSPVSPSVSGRGPDPDAHVAVNLSRYEG 419
Db 361 XPAHPVLSFLRDSMDLVSAFSLPLAPLSPTSPVSPSVSGRGPDPDAHVAVNLSRYEG 419

RESULT 15
AAE09202
ID AAE09202 standard; protein; 419 AA.
XX
XX AAE09202;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 generic protein variant (Arg371file).
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
XX ECDIIIA; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 1..340
XX /note= "Identical to N-terminal region of p185HER-2"
XX
XX Misc-difference 124 /note= "Represented as Agn in the parent sequence shown
XX in the specification"
XX
XX Misc-difference 125 /note= "Represented as Agn in the parent sequence shown
XX in the specification"
XX
XX Domain 341..419
XX /label= ECDIIIA variant
XX /note= "Extracellular domain IIIa variant"
XX
XX Misc-difference 342 /label= Unknown
XX
XX Misc-difference 345 /label= Unknown
XX
XX Misc-difference 346 /label= Unknown
XX
XX Misc-difference 356 /label= Unknown
XX
XX Misc-difference 358 /label= Unknown
XX
XX Misc-difference 361 /label= Unknown
XX
XX Misc-difference 371 /note= "p68HER-2 generic sequence (AAE09181) Arg
XX substituted with Ile"
XX
XX Misc-difference 376 /label= Unknown
XX
XX Misc-difference 394 /label= Unknown
XX
XX Misc-difference 404 /label= Unknown
XX
XX Misc-difference 413 /label= Unknown
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX

PI Clinton G, Henner WD, Evans A;
XX WPI; 2001-529934/58.
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX Example 12; Page; 61pp; English.
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 10⁸. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence. Note: The present sequence is not
XX shown in the specification but is derived from p68HER-2 generic sequence
XX (SEQ ID NO:2) shown in the sequence listing (AAE09181)
XX Sequence 419 AA;
XX Query Match 94.5%; Score 2161; DB 4; Length 419;
XX Best Local Similarity 95.2%; Pred. No. 8.4e-164;
XX Matches 399; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 1 MELAALCRWGLLLALLPGCAASTQVCTGDKMLRPLPASPETHLDMRLHYQCQVQGNL 60
Db 1 MELAALCRWGLLLALLPGCAASTQVCTGDKMLRPLPASPETHLDMRLHYQCQVQGNL 60
QY 61 EUTYLPNTASLSFLDIOIQVQGYLIANNQVQVPLQRLIRVGTQQLFEDNYVALVLDNG 120
Db 61 EUTYLPNTASLSFLDIOIQVQGYLIANNQVQVPLQRLIRVGTQQLFEDNYVALVLDNG 120
QY 121 DPLNNTPTVTGASPGSLRELQRLSLTEILKGGVLIQRNPQLCYQDTILWKDIFHRNQLA 180
Db 121 DPLNNTPTVTGASPGSLRELQRLSLTEILKGGVLIQRNPQLCYQDTILWKDIFHRNQLA 180
QY 181 LTLIDNRSRACHPCSPCKGRCWGESSEDCQSLTRTVACGACRCKGPIPTDCCHQC 240
Db 181 LTLIDNRSRACHPCSPCKGRCWGESSEDCQSLTRTVACGACRCKGPIPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFNHSIGICELHCPALVTYNTDIFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFNHSIGICELHCPALVTYNTDIFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPPLHQSVTAEDGTQRCCKSPKPCARGTHSLPRPAAVPLRMQP 360
Db 301 YNYLSTDVGSCTLVCPPLHQSVTAEDGTQRCCKSPKPCARGTHSLPRPAAVPLRMQP 360
QY 361 GPAHPVLSFLRPSWDLVSAFYSILPLAPLSPTSPISPSVSGRGPDPDAHVAVNLRYEG 419
Db 361 GPAHPVLSFLRPSWDLVSAFYSILPLAPLSPTSPISPSVSGRGPDPDAHVAVNLRYEG 419
RESULT 16
AAE09200
ID AAE09200 standard; protein; 419 AA.
XX AC AAE09200;
XX DT 15-NOV-2001 (first entry)
XX Human p68HER-2 generic protein variant (Arg357Cys).
XX DE HER-2, herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX KW solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
XX

KW ECDIIIA; variant.
XX OS Homo sapiens.
FH Key Location/Qualifiers
XX Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Domain 341..419
FT /label= ECDIIIA variant
FT /note= "Extracellular domain IIIa variant"
FT Misc-difference 342
FT /label= Unknown
FT Misc-difference 345
FT /label= Unknown
FT Misc-difference 346
FT /label= Unknown
FT Misc-difference 356
FT /label= Unknown
FT Misc-difference 357
FT /note= "p68HER-2 generic sequence (AAE09181) Arg
FT substituted with Cys"
FT Misc-difference 358
FT /label= Unknown
FT Misc-difference 361
FT /label= Unknown
FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown
XX WO200161356-A1.
XX 23-AUG-2001.
XX 16-FEB-2001; 2001WO-US005327.
XX 16-FEB-2000; 2000US-00506079.
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX Clinton G, Henner WD, Evans A;
XX WPI; 2001-529934/58.
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX Example 12; Page; 61pp; English.
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 10⁸. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence. Note: The present sequence is not
XX containing ECDIIIA variant sequence. Note: The present sequence is not

CC shown in the specification but is derived from p68HER-2 generic sequence
 CC (SEQ ID NO:2) shown in the sequence listing (AA09181)
 XX
 SQ Sequence 419 AA;

Query Match 94.5%; Score 2161; DB 4; Length 419;
 Best Local Similarity 95.2%; Pred. No. 8.4e-164;
 Matches 399; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLLALLPPGAASQVCTGDMKRLPASPETHLDMRLHYQGCVVQGNL 60
 DB 1 MELAALCRWGLLLALLPPGAASQVCTGDMKRLPASPETHLDMRLHYQGCVVQGNL 60
 QY 61 ELTYLPTNASLSFLQDIQEVQGVYLAHNOVROVPLQRLRIVRGTLQFEDNVALAVLDNG 120
 DB 61 ELTYLPTNASLSFLQDIQEVQGVYLAHNOVROVPLQRLRIVRGTLQFEDNVALAVLDNG 120
 QY 121 DPLNNTPTVTGASPGGLRQLRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNNOLA 180
 DB 121 DPLXXTPTVTGASPGGLRQLRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNNOLA 180
 QY 181 LTLIDTNRSRACHPCSPCKGSRGWSSSDQSLRTVCAGCARCKGPLEPTDCCHEQC 240
 DB 181 LTLIDTNRSRACHPCSPCKGSRGWSSSDQSLRTVCAGCARCKGPLEPTDCCHEQC 240
 QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
 DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
 QY 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSPKARGXHSXXPRPAAPVPLRMQP 360
 DB 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSPKARGXHSXXPRPAAPVPLRMQP 360
 QY 361 GPAHPVLSFLRPSWDLVSFAFSLPLAPLSPTSPVSVGRGPDPAHVAVNLRYEG 419
 DB 361 XPAHPVLSFLRPSWDXVSAFSLPLAPLPTSVKXISPVSVGRGXDPDAHVAVXLRYEG 419

RESULT 17
 AAY97240
 ID AAY97240 standard; protein; 420 AA.
 AC AAY97240;
 XX
 DT 04-DEC-2000 (first entry)
 DE
 DE Truncated HER-2, p68-HER-2.
 KW HER-2; erbB-2; oncogene; receptor-like tyrosine kinase; insertion;
 KW extracellular domain IIIa; antagonist; intron 8; C-terminal extension;
 KW truncated HER-2; p68; dimerization inhibitor; cytostatic.
 XX
 CS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 124
 FT Misc-difference 125 /note= "represented as Agn"
 FT Misc-difference 125 /note= "represented as Agn"
 FT Misc-difference 343
 FT Misc-difference 346 /note= "Preferably Ser"
 FT Misc-difference 346 /note= "Preferably Pro"
 FT Misc-difference 347 /note= "Preferably Leu"
 FT Misc-difference 357 /note= "Preferably Gln"
 FT Misc-difference 359 /note= "Preferably Leu"
 FT Misc-difference 362 /note= "changes from glycine"
 FT Misc-difference 377 /note= "Preferably Ile"

FT Misc-difference 395 /note= "Preferably Arg"
 FT Misc-difference 405 /note= "Preferably Leu"
 FT Misc-difference 414 /note= "Preferably Asn"
 XX
 XX WO200044403-A1.
 XX
 XX 03-AUG-2000.
 XX
 XX 20-JAN-2000; 2000WO-US001484.
 XX
 XX 20-JAN-1999; 99US-00234208.
 XX (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 XX Doherty JK, Clinton GM, Adelman JP;
 XX WPI; 2000-499287/44.
 XX
 XX Using polypeptides and antibodies that bind to the extracellular domain
 XX of the receptor-like tyrosine kinase HER-2 to treat solid tumors of the
 XX breast, lung, ovaries and colon.
 XX
 XX Claim 9; Page 39-40; 46pp; English.
 XX
 XX HER-2/neu (erbB-2) oncogene encodes a receptor-like tyrosine kinase. The
 XX extracellular domain of p185-HER-2 is proteolytically shed from breast
 XX carcinoma cells in culture and is found in serum of some cancer patients
 XX and may be a serum marker of metastatic breast cancer. An alternative HER
 XX -2 mRNA of 4.8 kb with a 274 bp insert (intron 8) has been identified.
 XX The retained intron is in-frame and encodes a 79 amino acid extension
 XX designated ECDIIa (the present sequence), which is inserted at residue
 XX 340 of p185-HER-2. The alternative mRNA predicts a truncated HER-2
 XX protein (approximately 68 kDa) that lacks the transmembrane and
 XX intracellular domains (see AAY97240). p68HER-2 specifically binds to p185
 XX HER-2 without activating HER-2. It could therefore block dimerization of
 XX p185-HER-2. The p68HER-2 polypeptide binds to a site on the ECD of HER-2
 XX that is different from the site of binding for Herceptin (RTM) (a
 XX marketed humanized monoclonal antibody that is used for the treatment of
 XX cancer and binds to the ECD of HER-2). The methods, compositions,
 XX polypeptides and antibodies are used to treat solid tumors such as
 XX breast cancer, small cell lung carcinoma, ovarian cancer and/or colon
 XX cancer, especially where over-expression of HER-2 is indicated
 XX
 XX SQ Sequence 420 AA;

Query Match 94.4%; Score 2158.5; DB 3; Length 420;
 Best Local Similarity 95.2%; Pred. No. 1.3e-163;
 Matches 400; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

QY 1 MELAALCRWGLLLALLPPGAASQVCTGDMKRLPASPETHLDMRLHYQGCVVQGNL 60
 DB 1 MELAALCRWGLLLALLPPGAASQVCTGDMKRLPASPETHLDMRLHYQGCVVQGNL 60
 QY 61 ELTYLPTNASLSFLQDIQEVQGVYLAHNOVROVPLQRLRIVRGTLQFEDNVALAVLDNG 120
 DB 61 ELTYLPTNASLSFLQDIQEVQGVYLAHNOVROVPLQRLRIVRGTLQFEDNVALAVLDNG 120
 QY 121 DPLNNTPTVTGASPGGLRQLRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNNOLA 180
 DB 121 DPLXXTPTVTGASPGGLRQLRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNNOLA 180
 QY 181 LTLIDTNRSRACHPCSPCKGSRGWSSSDQSLRTVCAGCARCKGPLEPTDCCHEQC 240
 DB 181 LTLIDTNRSRACHPCSPCKGSRGWSSSDQSLRTVCAGCARCKGPLEPTDCCHEQC 240
 QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
 DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
 QY 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSPKARGXHSXXPRPAAPVPLRMQP 359

Db 301 YNKLSTVGSCTVNCPLHNEQVTAEDGTQRCCKSKFCARVGHSHXAPRFAAVPVRXQ 360
QY 360 RCPAHPVLSFLRPSNDLVSAFYSLPLAPLSTSVIPSVSGRGFPDPAHVAVNLSRYEG 419
Db 361 XPAPHPVLSFLRPSNDLVSAFYSLPLAPLSTSVIPSVSGRGFPDPAHVAVNLSRYEG 420

RESULT 18
AAE09183
ID AAE09183 standard; protein; 419 AA.
XX AAE09183;
AC AAE09183;
XX 15-NOV-2001 (first entry)
DT Human p68HER-2 generic sequence #2.
DE HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
KW ECDIIIA.
XX
OS Homo sapiens.
XX
FH Key
FT Region
FT Location/Qualifiers
FT 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT /note= "Represented as Agn in the sequence shown in the
FT specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the sequence shown in the
FT specification"
FT Domain 341..419
FT /label= ECDIIIA
FT /note= "Extracellular domain IIIta"
FT Misc-difference 342
FT /label= Unknown
FT /note= "Encoded by WCC"
FT Misc-difference 345
FT /label= Unknown
FT /note= "Encoded by CYG"
FT Misc-difference 346
FT /label= Unknown
FT /note= "Encoded by CYC"
FT Misc-difference 356
FT /label= Unknown
FT /note= "Encoded by CWG"
FT Misc-difference 357
FT /label= Unknown
FT /note= "Encoded by YGC"
FT Misc-difference 358
FT /label= Unknown
FT /note= "Encoded by ATR"
FT Misc-difference 361
FT /label= Unknown
FT /note= "Encoded by GNC"
FT Misc-difference 371
FT /label= Unknown
FT /note= "Encoded by AKa"
FT Misc-difference 376
FT /label= Unknown
FT /note= "Encoded by MTA"
FT Misc-difference 389
FT /label= Unknown
FT /note= "Encoded by AGC"
FT Misc-difference 394
FT /label= Unknown
FT /note= "Encoded by CST"
FT Misc-difference 404
FT /label= Unknown
FT /note= "Encoded by CYG"
FT Misc-difference 413
FT /label= Unknown

/note= "Encoded by VAC"
WO200161356-A1.
23-AUG-2001.
15-FEB-2001; 2001WO-US005327.
16-FEB-2000; 2000US-00506079.
{UOR-} UNIV OREGON HEALTH SCI.
Clinton G, Hemmer WD, Evans A;
WPI; 2001-529934/58.
N-PSDB; RAD15852.
New polypeptide, which binds to the extracellular domain of HER-2 for the
treatment of hard tumors.
Claim 8; Page 57-58; 61pp; English.
The invention relates to novel HER-2 (herstatin-2) antagonist
particularly a polypeptide that binds to the extracellular domain (ECD)
of HER-2 at a site that is different from the binding site of humanised
antibody, Herceptin, at an affinity of at least 10⁸. The present
invention is based upon the initial discovery of an alternative HER-2
mRNA transcript with 274 bp insert of intron 8. The translation product
p68HER-2 which lacks the transmembrane and intracellular domains of
p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
nucleic acids encoding these are useful to treat, diagnose and identify
solid tumours. The present sequence is human p68HER-2 protein containing
ECDIIIA generic sequence
SQ Sequence 419 AA;
Query Match 94.3%; Score 2157; DB 4; Length 419;
Best Local Similarity 95.0%; Pred. No. 1.7e-163;
Matches 398; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 1 MELAALCRWGLLLALLPPGAASTQVCTGDMKRLPASPETHLDMRLHYQSCVVGQNL 60
Db 1 MELAALCRWGLLLALLPPGAASTQVCTGDMKRLPASPETHLDMRLHYQSCVVGQNL 60
QY 61 ELTYLPTNASISFLQDIQEVQGVLYAHNQVQVPLQRLRIVRGTLFEDNYALAVLNG 120
Db 61 ELTYLPTNASISFLQDIQEVQGVLYAHNQVQVPLQRLRIVRGTLFEDNYALAVLNG 120
QY 121 DPLNNTPTVTGASPGGLRHLQLSLTEILLKGGVLTORNPOLCYQDTILWKDIFHKNQLA 180
Db 121 DPLXNTPPTVTGASPGGLRHLQLSLTEILLKGGVLTORNPOLCYQDTILWKDIFHKNQLA 180
QY 181 LTLDITNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCAAGCARGKPLPTDCHEQC 240
Db 181 LTLDITNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCAAGCARGKPLPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMNPGRVTCGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMNPGRVTCGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNEQVTAEDGTQRCCKSKFCARVGHSHXAPRFAAVPVRXQ 360
Db 301 YNKLSTDVGSCTLVCPHNEQVTAEDGTQRCCKSKFCARVGHSHXAPRFAAVPVRXQ 360
QY 361 GPAHPVLSFLRPSNDLVSAFYSLPLAPLSTSVIPSVSGRGFPDPAHVAVNLSRYEG 419
Db 361 XPAHPVLSFLRPSNDLVSAFYSLPLAPLSTSVIPSVSGRGFPDPAHVAVNLSRYEG 419
RESULT 19

AAE09214
ID AAE09214 standard; protein; 419 AA.
AC AAE09214;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 generic protein variant 13.
XX
KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
ECDIIIA; variant.
XX
OS Homo sapiens.
XX
FE Key Location/Qualifiers
FT Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Misc-difference 124
FT /note= "Represented as Agn in the sequence shown in the
FT specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the sequence shown in the
FT specification"
FT Domain 341..419
FT /label= ECDIIIA variant
FT /note= "Extracellular domain IIIa"
FT Misc-difference 342
FT /label= Unknown
FT Misc-difference 345
FT /label= Unknown
FT Misc-difference 346
FT /label= Unknown
FT Misc-difference 356
FT /label= Unknown
FT Misc-difference 357
FT /note= "p68HER-2 generic sequence (AAE09183) Xaa
FT substituted with Cys"
FT Misc-difference 358
FT /label= Unknown
FT Misc-difference 361
FT /label= Unknown
FT Misc-difference 371
FT /label= Unknown
FT Misc-difference 376
FT /label= Unknown
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown
XX WO200161356-A1.
XX
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US005327.
XX
XX 16-FEB-2000; 2000US-00506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for the
XX treatment of hard tumors.
XX
XX Example 12; Page; 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist

CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10⁸. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIIA variant sequence. Note: The present sequence is not
CC shown in the specification but is derived from p68HER-2 generic sequence
CC (SEQ ID NO:13) shown in the sequence listing (AAE09183)
XX
XX Sequence 419 AA;
SQ
Query Match 94.2%; Score 2155; DB 4; Length 419;
Best Local Similarity 95.0%; Pred. No. 2.5e-163;
Matches 398; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 1 MELAALCRWGLLLALLPPGAASTQVCTGTDKMLPASPETHLDMLRLHYGCGVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASTQVCTGTDKMLPASPETHLDMLRLHYGCGVQGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQCYVLIHNVQVPLQLRLVRGTQLPEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQCYVLIHNVQVPLQLRLVRGTQLPEDNYALAVLDNG 120
QY 121 DPLANTTPTVGASPGGLRELQRLSLTEILKGVLIQNPOLCYQDTILWKDIFHKQNOLA 180
DB 121 DPLXXTTPTVGASPGGLRELQRLSLTEILKGVLIQNPOLCYQDTILWKDIFHKQNOLA 180
QY 181 LTLIDTNRSRACHPCSPCKGSRGWESSEDCQSLTRTVACAGGACRCKGPIPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPCKGSRGWESSEDCQSLTRTVACAGGACRCKGPIPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPLEHNVQVTAEDGTQRCCKSKPCARGTHSLPPPAVPVPLRMQP 360
DB 301 YNKJSTDVGSCTLVCPLEHNVQVTAEDGTQRCCKSKPCARGTHSLPPPAVPVPLRMQP 360
QY 361 GPAHPVLSFLRPSWDLVSAFYSLLPLAPLSPTSPVSPVSGRGPDPDAHVAVNLRYEG 419
DB 361 XPAHPVLSFLRPSWDLVSAFYSLLPLAPLPTSVXISPVSGRGPDPDAHVAVNLRYEG 419
RESULT 20
ID AAE09215
XX AAE09215 standard; protein; 419 AA.
XX
XX AAE09215;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 generic protein variant 14.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy; p68HER-2;
XX ECDIIIA; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 1..340
XX /note= "Identical to N-terminal region of p185HER-2"
XX Misc-difference 124
XX /note= "Represented as Agn in the sequence shown in the
XX specification"
XX

XX AC ABG70753;
 XX DT 28-NOV-2002 (first entry)
 XX DE Human HER2 receptor extracellular domain.
 XX KW Human; HER2; analyte; interfering substance; serum; HERCEPTIN; therapy;
 KW anti-HER2; antibody; plasma; HER2; ErbB2; ErbB2 receptor;
 KW extracellular domain; ECD; epitope; cancer.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= Signal_peptide
 XX US2002090662-A1.
 XX PD 11-JUL-2002.
 XX PF 01-AUG-2001; 2001US-00921161.
 XX PR 15-AUG-2000; 2000US-0225433P.
 XX PA (RALP/) RALPH P.
 XX PI Ralph P;
 XX DR WPI; 2002-597507/75.
 XX PT Determining an analyte in the presence of an interfering substance,
 PT comprises contacting a sample with an antibody recognizing analyte coated
 PT surface and interfering substance, and measuring unbound and bound
 PT analyte.
 XX PS Example 1; Fig 1; 16pp; English.
 XX CC The invention discloses a method for determining the amount of an analyte
 CC in a fluid sample in the presence of an interfering substance. The method
 CC comprises contacting a solid surface, dual-coated with an antibody
 CC recognising a free analyte and a second antibody recognising an
 CC interfering substance when bound to the analyte, with a fluid sample and
 CC then determining the total amount of free analyte and analyte bound to
 CC the interfering substance. The example in this specification discloses
 CC the quantitative determination of serum HERCEPTIN levels in patients
 CC undergoing HERCEPTIN therapy. The method accurately determines the amount
 CC of an anti-HER2 antibody in serum or plasma in the presence of an HER2
 CC (also referred to as ErbB2 and ErbB2 receptor, not defined; extracellular
 CC domain (ECD). The labelled secondary antibody recognises the anti-HER2
 CC antibody at an epitope different from that recognised by the first
 CC antibody, as well as different from that recognised by the HER2 ECD. The
 CC method is useful for accurately determining the amount of an antibody,
 CC anti-HER2, that is circulating in a fluid sample (serum or plasma)
 CC derived from a cancer patient undergoing anti-HER2 therapy, in the
 CC presence of an interfering substance (the ECD) of the HER2 oncogene. The
 CC method prevents loss of detection in the qualitative assays and an
 CC underestimation in the quantitative assays. The sequence presented is the
 CC human HER2 receptor extracellular domain
 XX SQ Sequence 645 AA;
 Query Match 82.1%; Score 1878; DB 5; Length 645;
 Best Local Similarity 83.0%; Pred. No. 5.3e-141;
 Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
 OY 1 MELALCRGLLALLPFGAASVCTGDMKLRLPASPEHLDMLRLHYQCGVQGNL 60
 DB 1 MELALCRGLLALLPFGAASVCTGDMKLRLPASPEHLDMLRLHYQCGVQGNL 60
 OY 61 ELTYLPTNASLFLQDIQEVQGVLIHQNVRQVPLQLRLIRVRGTLQFEDNYALAVLNG 120
 DB 61 ELTYLPTNASLFLQDIQEVQGVLIHQNVRQVPLQLRLIRVRGTLQFEDNYALAVLNG 120

OY 121 DPLNNTTPTVTCASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNOLA 180
 DB 121 DPLNNTTPTVTCASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNOLA 180
 OY 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDQSLTRTVCAAGCARCKGPLPTDCCHQC 240
 DB 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDQSLTRTVCAAGCARCKGPLPTDCCHQC 240
 OY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTFSSMPNPGRYTFGASCVTACP 300
 DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTFSSMPNPGRYTFGASCVTACP 300
 OY 301 YNYLSTDVGSCTLVCPLENQEVTAEDGTQRCCKSKPCAR-----GTHSLLPRAAVFVP 355
 DB 301 YNYLSTDVGSCTLVCPLENQEVTAEDGTQRCCKSKPCARVVCYGLGMEHLRREVAVTSAN 360
 OY 356 LRQPG--PAHPVLSFLRPSWDLVSIFYSLPLASLSTSVPI-----SPVSGRGED 405
 DB 361 IQEFAGCKKIFGSLAFIPESFGDPASNT---APLQPEQLQVFTLEEITGYLYISAWPD 417
 OY 406 --PDAHVAVNLRYEG 419
 DB 418 SLPDLVSFQNLQVIRG 433
 RESULT 24
 ADE71462
 ID ADE71462 standard; protein; 645 AA.
 AC ADE71462;
 XX 29-JAN-2004 (first entry)
 DT Human ErbB2.
 DE
 XX KW Cytostatic; ErbB2 antagonist; tyrosine kinase inhibitor; cancer; ErbB2;
 KW ErbB2 antibody; carcinoma; lymphoma; blastoma; sarcoma; liposarcoma;
 KW neuroendocrine tumour; mesothelioma; schwannoma; meningioma;
 KW adenocarcinoma; melanoma; leukaemia; lymphoid malignancy;
 KW squamous cell cancer; epithelial squamous cell cancer; lung cancer;
 KW small-cell lung cancer; non-small cell lung cancer;
 KW adenocarcinoma of the lung; squamous carcinoma of the lung;
 KW cancer of the peritoneum; hepatocellular cancer; gastric cancer;
 KW stomach cancer; gastrointestinal cancer; pancreatic cancer; glioblastoma;
 KW cervical cancer; ovarian cancer; liver cancer; bladder cancer; hepatoma;
 KW breast cancer; colon cancer; rectal cancer; colorectal cancer;
 KW endometrial carcinoma; uterine carcinoma; salivary gland carcinoma;
 KW kidney cancer; renal cancer; prostate cancer; vulval cancer;
 KW thyroid cancer; hepatic carcinoma; anal carcinoma; penile carcinoma;
 KW testicular cancer; oesophageal cancer; tumour of the biliary tract;
 KW head and neck cancer; human.
 XX OS Homo sapiens.
 XX US2003086924-A1.
 PD 08-MAY-2003.
 PF 10-OCT-2002; 2002US-00268501.
 PR 25-JUN-1999; 99US-0141316P.
 PR 23-JUN-2000; 2000US-00602812.
 XX (GETH) GENENTECH INC.
 XX Sliwowski MK;
 XX WPI; 2004-020226/02.
 PT Treating cancer (e.g. carcinoma, lymphoma or sarcoma) that expresses
 ErbB2 by administering to a patient an anti-ErbB2 antibody, and
 optionally an epidermal growth factor receptor-targeted drug or a

```
PT tyrosine kinase inhibitor.
XX Disclosure; SEQ ID NO 13; 56pp; English.
XX
XX The invention describes a method of treating cancer that expresses ErbB2
XX comprising administering to a patient an antibody that binds ErbB2.
XX Specifically claimed are antibodies that bind ErbB2, particularly
XX monoclonal antibody 2C4 or humanised 2C4, or monoclonal antibody 4D5 or
XX humanised 4D5. The methods are useful for treating cancer in a patient,
XX particularly a human. The cancer includes carcinoma, lymphoma, blastoma,
XX sarcoma, liposarcoma, neuroendocrine tumour, mesothelioma, schwannoma,
XX meningioma, adenocarcinoma, melanoma, leukaemia, lymphoid malignancy,
XX squamous cell cancer, epithelial squamous cell cancer, lung cancer, small
XX -cell lung cancer, non-small cell lung cancer, adenocarcinoma of the
XX lung, squamous carcinoma of the lung, cancer of the peritoneum,
XX hepatocellular cancer, gastric or stomach cancer, gastrointestinal
XX cancer, pancreatic cancer, glioblastoma, cervical cancer, ovarian cancer,
XX liver cancer, bladder cancer, hepatoma, breast cancer, colon cancer,
XX rectal cancer, colorectal cancer, endometrial or uterine carcinoma,
XX salivary gland carcinoma, kidney or renal cancer, prostate cancer, vulval
XX cancer, thyroid cancer, hepatic carcinoma, anal carcinoma, penile
XX carcinoma, testicular cancer, esophageal cancer, a tumour of the biliary
XX tract, or head and neck cancer. This is the amino acid sequence of human
XX ErbB2.
XX
XX Sequence 645 AA;
XX
Query Match      82.1%; Score 1878; DB 8; Length 645;
Best Local Similarity 83.0%; Pred. No. 5.3e-141;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
QY 1 MELAALCRWGLLLALLPPGAASQVCTGDMKLRLPASPETHLDMLRHLYQGCVVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASQVCTGDMKLRLPASPETHLDMLRHLYQGCVVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNVQVQLRIRVGTQIFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNVQVQLRIRVGTQIFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQNPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQNPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDNTRSRACHPCSPMCKSGRCWGESSEDQSLTRTVCAAGCARCKGKPLPTDCCHQC 240
DB 181 LTLIDNTRSRACHPCSPMCKSGRCWGESSEDQSLTRTVCAAGCARCKGKPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNEVTAEDGTQRCCKSKPCAR-----GTHSLPRPAAVFPV 355
DB 301 YNYLSTDVGSCTLVCPHNEVTAEDGTQRCCKSKPCAR-----GTHSLPRPAAVFPV 355
QY 356 LRMQPG--PAHPVLSFLRPSNDLVSAFYSPLAPLSPTSVPI-----SPVSVGRGPD 405
DB 356 LRMQPG--PAHPVLSFLRPSNDLVSAFYSPLAPLSPTSVPI-----SPVSVGRGPD 405
QY 361 IQEPAGCKKIFGSLAPLPESFDGDPASNT---APLQPEQLQVFTLEETIGLYISAWPD 417
DB 361 IQEPAGCKKIFGSLAPLPESFDGDPASNT---APLQPEQLQVFTLEETIGLYISAWPD 417
QY 406 --PDAHVAVNLRYEG 419
DB 418 SLPDLVSFQNLQVIRG 433
RESULT 25
ID AAB21200 standard; protein; 653 AA.
XX
XX AAB21200;
XX
XX 12-JAN-2001 (first entry)
XX
XX Extracellular HER-2/neu protein.
XX
```

```
XX HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine; breast cancer;
XX prostate cancer; ovarian cancer; lung cancer; colon cancer.
XX
XX Unidentified.
XX
XX WO200044899-A1.
XX
XX 03-AUG-2000.
XX
XX 28-JAN-2000; 2000WO-US002164.
XX
XX 29-JAN-1999; 99US-0117976P.
XX
XX (CORI-) CORIXA CORP.
XX (SMIK) SMITHLINE BEECHAM.
XX
XX Cheever MA, Gheysen D;
XX
XX WPI; 2000-505976/45.
XX
XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins
XX useful for vaccinating against breast, ovarian, colon, lung and prostate
XX cancers.
XX
XX Claim 2; Fig 9; 128pp; English.
XX
XX The present sequence is the extracellular HER-2/neu protein. HER-2/neu is
XX a member of the tyrosine kinase family of receptor-like glycoproteins and
XX shows homology to the epidermal growth factor receptor (EGFR). It
XX probably plays a part in cell growth and/or differentiation. The HER-
XX 2/neu gene is an oncogene. An HER-2/neu fusion protein comprising a HER-
XX 2/neu extracellular domain fused to a HER-2/neu phosphorylation domain
XX may be used to treat or prevent cancer by eliciting or enhancing an
XX immune response to the HER-2/neu protein. It may be used to treat
XX malignancies such as breast, ovarian, colon, lung and prostate cancers,
XX and may be used as an antigen to vaccinate against these neoplasias
XX
XX Sequence 653 AA;
XX
Query Match      82.1%; Score 1878; DB 3; Length 653;
Best Local Similarity 83.0%; Pred. No. 5.4e-141;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
QY 1 MELAALCRWGLLLALLPPGAASQVCTGDMKLRLPASPETHLDMLRHLYQGCVVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASQVCTGDMKLRLPASPETHLDMLRHLYQGCVVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNVQVQLRIRVGTQIFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNVQVQLRIRVGTQIFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQNPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQNPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDNTRSRACHPCSPMCKSGRCWGESSEDQSLTRTVCAAGCARCKGKPLPTDCCHQC 240
DB 181 LTLIDNTRSRACHPCSPMCKSGRCWGESSEDQSLTRTVCAAGCARCKGKPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNEVTAEDGTQRCCKSKPCAR-----GTHSLPRPAAVFPV 355
DB 301 YNYLSTDVGSCTLVCPHNEVTAEDGTQRCCKSKPCAR-----GTHSLPRPAAVFPV 355
QY 356 LRMQPG--PAHPVLSFLRPSNDLVSAFYSPLAPLSPTSVPI-----SPVSVGRGPD 405
DB 356 LRMQPG--PAHPVLSFLRPSNDLVSAFYSPLAPLSPTSVPI-----SPVSVGRGPD 405
QY 361 IQEPAGCKKIFGSLAPLPESFDGDPASNT---APLQPEQLQVFTLEETIGLYISAWPD 417
DB 361 IQEPAGCKKIFGSLAPLPESFDGDPASNT---APLQPEQLQVFTLEETIGLYISAWPD 417
QY 406 --PDAHVAVNLRYEG 419
```


CC receptor-like glycoproteins and shows homology to the epidermal growth
 CC factor receptor (EGFR). It probably plays a part in cell growth and/or
 CC differentiation. The HER-2/neu gene is an oncogene. HER-2/neu fusion
 CC proteins may be used to treat or prevent cancer by eliciting or enhancing
 CC an immune response to the HER-2/neu protein. They may be used to treat
 CC malignancies such as breast, ovarian, colon, lung and prostate cancers,
 CC and may be used as an antigen to vaccinate against these neoplasias
 XX
 SQ Sequence 712 AA;

Query Match 82.1%; Score 1878; DB 3; Length 712;
 Best Local Similarity 83.0%; Pred. No. 6e-141;
 Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
 QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLMRLHLYQGCVVQGNL 60
 DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLMRLHLYQGCVVQGNL 60
 QY 61 ELTYLPTNASLSFLQDIOEVQGVLIHNNQVQVPLQRLIRIVRGTLFEDNYALAVLDNG 120
 DB 61 ELTYLPTNASLSFLQDIOEVQGVLIHNNQVQVPLQRLIRIVRGTLFEDNYALAVLDNG 120
 QY 121 DPLNNTTPTVGTASPGGLREQLRSLEILKGGVLIQNPOLCYQDTILWKDIFHKNNQLA 180
 DB 121 DPLNNTTPTVGTASPGGLREQLRSLEILKGGVLIQNPOLCYQDTILWKDIFHKNNQLA 180
 QY 181 LTLIDTNRSRACHPCSPMCKSGRCWGESSEDCOSLRTVTCAGGCARCKGPLETDCCHEQC 240
 DB 181 LTLIDTNRSRACHPCSPMCKSGRCWGESSEDCOSLRTVTCAGGCARCKGPLETDCCHEQC 240
 QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRYTFGASCVTACP 300
 DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRYTFGASCVTACP 300
 QY 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSKPCAR-----GTHSLPRPAVPVP 355
 DB 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSKPCAR-----GTHSLPRPAVPVP 355
 QY 356 LRWQPG--PAHPVLSFLRPSWDLVSAPYSLPLAPLSPTSVPI-----SPVSVGRGPD 405
 DB 361 IQEFAGCKKIFGSLAFPLPESFGDPASNT---APLOPEQLQVFTLEETGTYLISAWPD 417
 QY 406 --PDAHVNLSRYEG 419
 DB 418 SLEDLSVFQNLQVIRG 433

RESULT 28

AAM51149
 ID AAM51149 standard; protein; 712 AA.
 XX
 AC AAM51149;
 XX
 DT 17-JUN-2002 (first entry)
 XX
 DS Her-2/neu extracellular domain-delta-phosphorylation domain fusion.
 XX
 XW Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;
 XW tyrosine kinase; receptor; c-erbB2; gene therapy.
 XX
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH Domain 1..653
 FT /note= "extracellular domain"
 FT 654..712
 FT /note= "phosphorylation domain fragment"
 XX
 PN WO200212341-A2.
 XX
 PD 14-FEB-2002.
 XX
 PP 03-AUG-2001; 2001WO-US024283.

XX

03-AUG-2000; 2000US-00632507.

PR

{CORI-} CORIXA CORP.

PA

{SMIK } SMITHKLINE BEECHAM BIOLOGICALS.

XX

Cheever MA, Gheysen D;

PI

WPI; 2002-241743/29.

DR

XX Her-2/neu fusion protein for treating or preventing cancer by eliciting
 PT or enhancing an immune response to the protein, has Her-2/neu
 PT extracellular domain fused to Her-2/neu intracellular or phosphorylation
 PT domain.

XX

Claim 37; Fig 13; 141pp; English.

XX

CC The present sequence is that of a fusion protein between the
 CC extracellular domain and a fragment (DeltaPD) of the phosphorylation
 CC domain of human Her-2/neu (see AAM51143), an oncogenic self-protein and
 CC target for anti-cancer vaccines. The fusion protein can be obtained by
 CC recombinant DNA methods. Her-2/neu overexpression correlates with a poor
 CC prognosis in breast and ovarian cancers. The invention provides Her-2/neu
 CC fusion proteins, nucleic acids encoding them, viral vectors, and vaccines
 CC comprising the fusion proteins or nucleic acid molecules. In preferred
 CC fusion proteins, the extracellular domain of Her-2/neu is fused to a Her-
 CC 2/neu intracellular domain or phosphorylation domain (or its DeltaPD
 CC fragment). An immune response to Her-2/neu protein is elicited or
 CC enhanced by administering the fusion protein in the form of a vaccine, or
 CC by transfecting cells of an animal ex vivo with a nucleic acid encoding
 CC the fusion protein, and delivering the transfected cells to the animal.
 CC The fusion proteins, nucleic acids, and isolated specific T-cells are
 CC useful for inhibiting the development of a cancer, especially breast,
 CC ovarian, colon, lung or prostate cancer in a patient. T cells that
 CC specifically react with a Her-2/neu fusion protein can be used to remove
 CC tumour cells from a sample in order to inhibit the development of cancer
 CC in a patient

XX

SQ Sequence 712 AA;

Query Match 82.1%; Score 1878; DB 5; Length 712;
 Best Local Similarity 83.0%; Pred. No. 6e-141;
 Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY

1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLMRLHLYQGCVVQGNL 60

DB

1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLMRLHLYQGCVVQGNL 60

QY

61 ELTYLPTNASLSFLQDIOEVQGVLIHNNQVQVPLQRLIRIVRGTLFEDNYALAVLDNG 120

DB

61 ELTYLPTNASLSFLQDIOEVQGVLIHNNQVQVPLQRLIRIVRGTLFEDNYALAVLDNG 120

QY

121 DPLNNTTPTVGTASPGGLREQLRSLEILKGGVLIQNPOLCYQDTILWKDIFHKNNQLA 180

DB

121 DPLNNTTPTVGTASPGGLREQLRSLEILKGGVLIQNPOLCYQDTILWKDIFHKNNQLA 180

QY

181 LTLIDTNRSRACHPCSPMCKSGRCWGESSEDCOSLRTVTCAGGCARCKGPLETDCCHEQC 240

DB

181 LTLIDTNRSRACHPCSPMCKSGRCWGESSEDCOSLRTVTCAGGCARCKGPLETDCCHEQC 240

QY

241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRYTFGASCVTACP 300

DB

241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRYTFGASCVTACP 300

QY

301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSKPCAR-----GTHSLPRPAVPVP 355

DB

301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSKPCAR-----GTHSLPRPAVPVP 355

QY

356 LRWQPG--PAHPVLSFLRPSWDLVSAPYSLPLAPLSPTSVPI-----SPVSVGRGPD 405

DB

361 IQEFAGCKKIFGSLAFPLPESFGDPASNT---APLOPEQLQVFTLEETGTYLISAWPD 417

QY

406 --PDAHVNLSRYEG 419


```

Db      418 SLPDLSVFQNLQVIRG 433
|| | | | |
RESULT 29
AAW19764
ID AAW19764 standard; protein; 782 AA.
XX
AC AAW19764;
XX
DT 17-SEP-1997 (first entry)
XX
DE Her2-GM-CSF immunostimulant fusion protein.
XX
KW Her2-GM-CSF; granulocyte macrophage colony stimulating factor;
KW growth factor receptor; oncogene; immunostimulant; cancer; therapy.
XX
OS Homo sapiens.
FH
FT Key Location/Qualifiers
FT Protein 1..653
FT Peptide /label= Her2
FT /label= Linker
FT /label= 655
FT /label= 656..782
FT /label= GM-CSF
XX
PN WC9724438-Al.
XX
PD 10-JUL-1997.
XX
PF 23-DEC-1996; 96WO-US020241.
XX
PR 28-DEC-1995; 95US-00579823.
XX
PA (ACT1-) ACTIVATED CELL THERAPY INC.
XX
PI Laus R, Ruegg CL, Wu H;
XX
PI WPI; 1997-363674/33.
XX
DR N-PSDB; AAT72725.
XX
PT Potent APC that activates T-cells to give multivalent cellular immune
PT response - can also induce a cytotoxic T-cell response in a vertebrate
PT subject.
XX
PS Disclosure; Fig 8; 45pp; English.
XX
CC A fusion protein (AAW19764) comprises Her2 (a growth factor receptor that
CC is overexpressed in breast, ovarian and other cancer cells) and
CC granulocyte-macrophage colony stimulating factor (GM-CSF). It is the
CC expression product of a nucleic acid molecule (AAT72725) prep. by PCR
CC amplification of Her2 cDNA from a breast cancer cell line and fusion to
CC GM-CSF cDNA. Fusion expression vectors can be used to transfect mammalian
CC and insect cells. The Her2-GM-CSF fusion protein is used to generate anti
CC -Her2 immunity. Tumour cells are eliminated by cytotoxic T lymphocytes
CC activated in vivo or in vitro by exposure to antigen-presenting cells
CC exposed to the fusion protein
XX
SQ Sequence 782 AA;

Query Match 82.1%; Score 1878; DB 2; Length 782;
Best Local Similarity 83.0%; Pred No. 6.7e-141;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRGLLALLPGAASTQVCTGDMKRLPASPETHLDMRLHYQCCVQGNL 60
DB 1 MELAALCRGLLALLPGAASTQVCTGDMKRLPASPETHLDMRLHYQCCVQGNL 60
QY 61 ELTYLPTNASTSLFQDIQVQGVLIHNNQVQVPLQRLIRVGTQFEDNYALAVLNG 120
DB 61 ELTYLPTNASTSLFQDIQVQGVLIHNNQVQVPLQRLIRVGTQFEDNYALAVLNG 120

QY 121 DPLNNTTPTVGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHNNOLA 180
DB 121 DPLNNTTPTVGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHNNOLA 180
QY 181 LTLIDTNRSRACHPCSPMCKGSRGWCSSBDCQSLTRITVCAGGCARCKGPLPTDCCHEQC 240
DB 181 LTLIDTNRSRACHPCSPMCKGSRGWCSSBDCQSLTRITVCAGGCARCKGPLPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTNTDTFFSMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTNTDTFFSMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCITVCPILHNOEVTAEADGTQCEKCKPCAR-----GTHSLLPRAAIVPVP 355
DB 301 YNYLSTDVGSCITVCPILHNOEVTAEADGTQCEKCKPCARCYCLGMEHLREVEAVTISAN 360
QY 356 LRMPQG--PAHPVLSFLRPSNDLWSAFYSLPLAPLSPTSVPI-----SPVSVGRGPD 405
DB 361 IQEPAGCKKIFGSLAFTPESPFDGDPASNT---APLQPEQLQVFETLBEITGYLYISAWPD 417
QY 406 --PDAHVAVNLSRYEG 419
DB 418 SLPDLSVFQNLQVIRG 433

RESULT 30
AAB21203
ID AAB21203 standard; protein; 919 AA.
XX
AC AAB21203;
XX
DT 12-JAN-2001 (first entry)
XX
DE Human HER-2/neu fusion protein.
XX
KW Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
KW colon cancer; fusion protein.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200044899-Al.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-US002164.
XX
PR 29-JAN-1999; 99US-0117976P.
XX
PA (CORI-) CORIXA CORP.
PA (SMIK) SMITHKLINE BEECHAM.
XX
PI Cheever WA, Gheysen D;
XX
PI WPI; 2000-505976/45.
XX
PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and prostate
PT cancers.
XX
PS Claim 2; Fig 12; 128pp; English.
XX
CC The present sequence is a fusion protein comprising the extracellular
CC domain and the phosphorylation domain of the human HER-2/neu protein. HER
CC -2/neu is a member of the tyrosine kinase family of receptor-like
CC glycoproteins and shows homology to the epidermal growth factor receptor
CC (EGFR). It probably plays a part in cell growth and/or differentiation.
CC The HER-2/neu gene is an oncogene. HER-2/neu fusion proteins may be used
CC to treat or prevent cancer by eliciting or enhancing an immune response
CC to the HER-2/neu protein. They may be used to treat malignancies such as
CC breast, ovarian, colon, lung and prostate cancers, and may be used as an
CC antigen to vaccinate against these neoplasias

```

XX
SQ Sequence 919 AA;

Query Match 82.1%; Score 1879; DB 3; Length 919;
Best Local Similarity 83.0%; Pred. No. 8.2e-141;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLALPPGAASQVCTGTDMLRLPASPTHLDMLRHLVQGCQVVGQNL 60
DB 1 MELAALCRWGLLALPPGAASQVCTGTDMLRLPASPTHLDMLRHLVQGCQVVGQNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVYLIAHNQVROVPLQRLRIVRGTOLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVYLIAHNQVROVPLQRLRIVRGTOLFEDNYALAVLDNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILAKDIFHKNQOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILAKDIFHKNQOLA 180
QY 181 LTLIDTNRSRACHPCSPMCKGSRGWGESSEDCQSLTRTVCAAGCARCKGPLPTDCCHEQC 240
DB 181 LTLIDTNRSRACHPCSPMCKGSRGWGESSEDCQSLTRTVCAAGCARCKGPLPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACLHPNHSIGICELHCPALVTYNTDTFESMENPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLHPNHSIGICELHCPALVTYNTDTFESMENPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQCEKCKSPCAR-----GTHSLLPRAAPVP 355
DB 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQCEKCKSPCARCYGLGMEHLREVRVTSAN 360
QY 356 LRMQPG--PAHPVLSFLRPSWDLVSAPYSLPLAPLSPTSVP-----SPVSVGRGPD 405
DB 361 IQEFAGCKKIFGSLAFLPESFDGDPASNT---APLOPEQLQVPETLEEITGYLYISAWPD 417
QY 406 --PDHVVAVNLSRYEG 419
DB 418 SLPDLSVFQNLQVIRG 433

Search completed: July 4, 2004, 04:18:51
Job time : 91.6606 secs

OM protein - protein search, using sw model
Run on: July 4, 2004, 04:18:59 ; Search time 27.7651 seconds
(without alignments)
779.083 Million cell updates/sec

Title: US-09-506-079H-12
Perfect score: 2287
Sequence: 1 MELAALCRWGLLALLPPGA.....VGRGPDPAHVAVNLRYEG 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 90 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2287	100.0	419	4	US-09-630-155-2
2	1878	82.1	782	2	US-09-146-283-4
3	1878	82.1	782	3	US-08-579-823A-4
4	1878	82.1	782	3	US-09-344-195-4
5	1878	82.1	1255	1	US-08-467-083-68
6	1878	82.1	1255	1	US-08-414-417B-68
7	1878	82.1	1255	2	US-08-484-438-8
8	1878	82.1	1255	2	US-08-486-348A-68
9	1878	82.1	1255	2	US-08-625-101-2
10	1878	82.1	1255	2	US-08-458-545B-68
11	1878	82.1	1255	2	US-08-356-786-2
12	1878	82.1	1255	3	US-08-486-680B-68
13	1878	82.1	1255	4	US-09-527-487-2
14	1878	82.1	1255	4	US-08-811-115-3
15	1878	82.1	1255	4	US-09-354-533-68
16	1769	77.4	624	3	US-08-422-108-1
17	1769	77.4	624	4	US-08-422-734-1
18	860	37.6	166	4	US-08-648-067A-1
19	793	34.7	644	1	US-08-336-708A-9
20	793	34.7	1210	2	US-08-484-438-7
21	793	34.7	1210	2	US-08-475-035-4
22	775	33.9	911	2	US-08-494-438-10
23	775	33.9	1058	2	US-08-484-438-4
24	775	33.9	1308	2	US-08-484-438-2
25	773	33.8	478	4	US-09-570-454-2
26	773	33.8	478	4	US-09-867-521-2
27	735.5	32.2	1342	1	US-07-978-895-4

ALIGNMENTS

RESULT 1
US-09-630-155-2
; Sequence 2, Application US/09630155
; Patent No. 641130
; GENERAL INFORMATION:

28	735.5	32.2	1342	2	US-08-484-438-9	Sequence 9, Appli
29	735.5	32.2	1342	2	US-08-473-119-4	Sequence 4, Appli
30	735.5	32.2	1342	2	US-08-475-352-4	Sequence 4, Appli
31	735.5	32.2	1342	4	US-09-170-699-4	Sequence 4, Appli
32	734	32.1	1343	6	5183884-4	Patent No. 5183884
33	493	21.6	97	1	US-08-421-356-3	Sequence 3, Appli
34	493	21.6	97	4	US-09-046-783-3	Sequence 3, Appli
35	418	18.3	79	4	US-09-630-155-1	Sequence 1, Appli
36	264.5	11.6	1382	2	US-08-737-715-2	Sequence 2, Appli
37	264.5	11.6	1382	4	US-09-457-040B-7	Sequence 7, Appli
38	257.5	11.3	1367	3	US-08-746-559A-4	Sequence 4, Appli
39	257.5	11.3	1367	2	US-08-249-687C-2	Sequence 2, Appli
40	257.5	11.3	1367	2	US-08-625-819-2	Sequence 2, Appli
41	257.5	11.3	1367	3	US-08-746-559A-2	Sequence 2, Appli
42	257.5	11.3	1367	4	US-08-864-641B-18	Sequence 18, Appli
43	257.5	11.3	1367	4	US-09-343-551-2	Sequence 2, Appli
44	241.5	10.6	486	3	US-08-746-559A-5	Sequence 5, Appli
45	210.5	9.2	383	3	US-08-857-076-105	Sequence 105, App
46	203	8.9	1724	3	US-08-857-076-12	Sequence 12, Appl
47	196	8.6	366	3	US-08-857-076-103	Sequence 103, App
48	184.5	8.1	370	3	US-08-857-076-104	Sequence 104, App
49	175	7.7	32	4	US-09-648-067A-2	Sequence 2, Appli
50	147.5	6.4	381	3	US-08-857-076-106	Sequence 106, App
51	142	6.2	1940	2	US-08-844-271-30	Sequence 30, Appl
52	142	6.2	1940	4	US-09-077-955-34	Sequence 34, Appl
53	131.5	5.7	799	2	US-08-525-940-23	Sequence 23, Appl
54	131.5	5.7	799	2	US-08-976-838-23	Sequence 23, Appl
55	131.5	5.7	881	2	US-08-525-940-21	Sequence 21, Appl
56	131.5	5.7	881	2	US-08-976-838-21	Sequence 21, Appl
57	131.5	5.7	935	2	US-08-525-940-18	Sequence 18, Appl
58	131.5	5.7	935	2	US-08-976-838-18	Sequence 18, Appl
59	131.5	5.7	935	4	US-09-214-555B-7	Sequence 7, Appli
60	130.5	5.7	833	4	US-09-013-895A-5	Sequence 5, Appli
61	130.5	5.7	833	4	US-09-448-868-5	Sequence 2, Appli
62	129.5	5.7	915	4	US-09-214-555B-2	Sequence 2, Appli
63	128.5	5.6	420	4	US-09-907-794B-109	Sequence 109, App
64	128.5	5.6	420	4	US-09-305-125A-109	Sequence 109, App
65	128.5	5.6	420	4	US-09-302-775A-109	Sequence 109, App
66	126	5.5	288	1	US-08-368-852-15	Sequence 15, Appl
67	124	5.4	288	2	US-08-525-940-15	Sequence 15, Appl
68	124	5.4	288	2	US-08-976-838-15	Sequence 15, Appl
69	124	5.4	3075	2	US-08-460-309-5	Sequence 5, Appli
70	124	5.4	3075	2	US-08-125-077-5	Sequence 5, Appli
71	123	5.4	242	4	US-09-312-283C-393	Sequence 393, App
72	123	5.4	417	3	US-08-815-469-4	Sequence 4, Appli
73	123	5.4	417	3	US-09-153-927-2	Sequence 2, Appli
74	123	5.4	417	4	US-09-565-918-5	Sequence 5, Appli
75	123	5.4	417	4	US-08-928-069-10	Sequence 10, Appl
76	123	5.4	417	4	US-08-828-683A-6	Sequence 6, Appli
77	123	5.4	428	3	US-08-815-469-2	Sequence 2, Appli
78	118.5	5.2	3655	4	US-09-845-583A-2	Sequence 2, Appli
79	117	5.1	459	1	US-08-313-288B-15	Sequence 15, Appl
80	116.5	5.1	969	2	US-08-447-642-2	Sequence 2, Appli
81	116.5	5.1	969	2	US-08-447-642-2	Sequence 2, Appli
82	116.5	5.1	969	3	US-09-236-503-2	Sequence 2, Appli
83	116.5	5.1	969	5	PCT-US93-02147A-2	Sequence 2, Appli
84	115	5.0	1345	2	US-08-977-767-3	Sequence 3, Appli
85	115	5.0	221	4	US-09-728-884-1	Sequence 1, Appli
86	115	5.0	3084	4	US-09-562-702A-12	Sequence 12, Appl
87	115	5.0	3106	4	US-09-562-702A-10	Sequence 10, Appl
88	114.5	5.0	484	2	US-08-252-493C-9	Sequence 9, Appli
89	114.5	5.0	484	3	US-09-276-197-9	Sequence 9, Appli
90	114	5.0	248	4	US-09-252-991A-29249	Sequence 29249, A

APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSES: DAVIS WRIGHT TREMAINE LLP
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: PC compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/630,155
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Davison, Barry L.
REGISTRATION NUMBER: 47,309
REFERENCE/DOCKET NUMBER: 49321-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 628-7621
TELEFAX: 206 628-7699
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: polypeptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-630-155-2

Query Match 100.0%; Score 2287; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 3.2e-192;
Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPEHLDMLRHLYQGCVVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPEHLDMLRHLYQGCVVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVYLI AHNQVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVYLI AHNQVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
QY 121 DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFKKNOLA 180
DB 121 DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFKKNOLA 180
QY 181 LTLIDTNRSRACHPCSPCKGRCWGSSEDCQSLTRTVCAAGCARCKGPLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPCKGRCWGSSEDCQSLTRTVCAAGCARCKGPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNEVTAEDGTQRCCKSKPCARCTRTVCAAGCARCKGPLPTDCCHQC 360
DB 301 YNYLSTDVGSCTLVCPHNEVTAEDGTQRCCKSKPCARCTRTVCAAGCARCKGPLPTDCCHQC 360
QY 361 GPAHPVLSTFLRPSWDLVSAFYSLPLAPLSTPSVIPSPVSVGRGDPDPAHVAVNLSRYEG 419
DB 361 GPAHPVLSTFLRPSWDLVSAFYSLPLAPLSTPSVIPSPVSVGRGDPDPAHVAVNLSRYEG 419

RESULT 2
US-09-146-283-4
Sequence 4, Application US/09146283
Patent No. 5976546

GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Ruegg, Curtis L.
APPLICANT: Wu, Hongyu
TITLE OF INVENTION: Immunostimulatory Compositions
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSES: Dehlinger & Associates
STREET: 350 Cambridge Ave. Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/146,283
FILING DATE: 03-SEPT-1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702
REFERENCE/DOCKET NUMBER: 7636-0010.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 782 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-09-146-283-4

Query Match 82.1%; Score 1878; DB 2; Length 782;
Best Local Similarity 83.0%; Pred. No. 4.9e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPEHLDMLRHLYQGCVVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPEHLDMLRHLYQGCVVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVYLI AHNQVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVYLI AHNQVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
QY 121 DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFKKNOLA 180
DB 121 DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFKKNOLA 180
QY 181 LTLIDTNRSRACHPCSPCKGRCWGSSEDCQSLTRTVCAAGCARCKGPLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPCKGRCWGSSEDCQSLTRTVCAAGCARCKGPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNEVTAEDGTQRCCKSKPCARCTRTVCAAGCARCKGPLPTDCCHQC 355
DB 301 YNYLSTDVGSCTLVCPHNEVTAEDGTQRCCKSKPCARCTRTVCAAGCARCKGPLPTDCCHQC 355
QY 356 LRMQPG--PAHPVLSTFLRPSWDLVSAFYSLPLAPLSTPSVIPSPVSVGRGDPD 405
DB 361 IQEFAGCKKIFGSLAPLPSFDGDPASNT---APLQPEQLQVFETLEEITGLYIAGWPD 417

QY 406 --PDHVAVNLRYEG 419
DB 418 SLPDLVFNQLQVIRG 433

RESULT 3
US-08-579-823A-4
Sequence 4, Application US/08579823A
Patent No. 6080409
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Wu, Hongyu
TITLE OF INVENTION: Immunostimulatory Composition and Method
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave. Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/579,823A
FILING DATE: 03-DEC-1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702
REFERENCE/DOCKET NUMBER: 7636-0010
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0960
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 782 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-08-579-823A-4

Query Match 82.1%; Score 1878; DB 3; Length 782;
Best Local Similarity 83.0%; Pred. No. 4.9e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLLALLPPGAASVCTGTDMLKRLPASPETHLDMRLHYQGCQVVGQNL 60
DB 1 MELAALCRWGLLLALLPPGAASVCTGTDMLKRLPASPETHLDMRLHYQGCQVVGQNL 60

QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIHQNQVQVPLQRLIRVAGTQFEDNYALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHQNQVQVPLQRLIRVAGTQFEDNYALAVLNG 120

QY 121 DPLNNITPVTVGASPGGLREQLRLSLTEILKGGVLIQNPOLCYQDITLAKDIFHKQNQLA 180
DB 121 DPLNNITPVTVGASPGGLREQLRLSLTEILKGGVLIQNPOLCYQDITLAKDIFHKQNQLA 180

QY 181 LTLLIDNRSRACHPCSPMCKSGRCWGESSEDCQSLRTITVCAGGCARCKGPLETDCCHEQC 240
DB 181 LTLLIDNRSRACHPCSPMCKSGRCWGESSEDCQSLRTITVCAGGCARCKGPLETDCCHEQC 240

QY 241 AAGCTGPKSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTLVCPFLHNQEVTAEDGTQRCCKSKPCAR-----GTHSLLRPAAPVVP 355
DB 301 YNYLSTDVGSCTLVCPFLHNQEVTAEDGTQRCCKSKPCARCYGLGMEHLREVRVTSAN 360

QY 356 LRMPQG--PAHPVLSFLSPSWDLVSAPYSLPLAPLSPTSVP-----SPVSVGRGPD 405
DB 361 IQEPAGCKKIFGSLAPLPESEFDGDPASNT--APLQPEQLQVFFETLEETGYLYISANPD 417

QY 406 --PDHVAVNLRYEG 419
DB 418 SLPDLVFNQLQVIRG 433

RESULT 4
US-09-344-195-4
Sequence 4, Application US/09344195
Patent No. 6210662
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Wu, Hongyu
TITLE OF INVENTION: Immunostimulatory Compositions
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave. Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/344,195
FILING DATE: 24-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/146,283
FILING DATE: 03-SEPT-1998
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702
REFERENCE/DOCKET NUMBER: 7636-0010.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 782 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-09-344-195-4

Query Match 82.1%; Score 1878; DB 3; Length 782;
Best Local Similarity 83.0%; Pred. No. 4.9e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLLALLPPGAASVCTGTDMLKRLPASPETHLDMRLHYQGCQVVGQNL 60
DB 1 MELAALCRWGLLLALLPPGAASVCTGTDMLKRLPASPETHLDMRLHYQGCQVVGQNL 60

QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIHQNQVQVPLQRLIRVAGTQFEDNYALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHQNQVQVPLQRLIRVAGTQFEDNYALAVLNG 120

Db 61 ELTYLPTNASLSFLQDIQEVQGVLLIAHNOVRQVPLQRLRIVRGTQFDPEDNYALAVLDNG 120
Qy 121 DPLNNTTPTVGTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
Db 121 DPLNNTTPTVGTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
Qy 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGGCARCKGPLPTDCCHQC 240
Db 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGGCARCKGPLPTDCCHQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPGRYTTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPGRYTTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR-----GTHSLLPRAAIVPVP 355
Db 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARCYGLGMEHLREVAVTSAN 360
Qy 356 LRMQPG--PAHPVLSFLRPSNDLVSAFYSLPLAPLSPTSVP-----SPVSVGRGPD 405
Db 361 IQBFAGCKKIFGSLAFPLSPFSGDPAASNT---APLQPEQLQVFTLEITGYLYISAWPD 417
Qy 406 --PDAHVAVNLSRYEG 419
Db 418 SLFDLSVFQNLQVIRG 433

RESULT 5
US-08-467-083-68
; Sequence 68, Application US/08467083
; Patent No. 5726023
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,083
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/414,417
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 622-6031
; TELEX: 3723836 SEEDANBERRY
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-467-083-68

Query Match 82.1%; Score 1878; DB 1; Length 1255;

Best Local Similarity 83.0%; Pred. No. 9.1e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
Qy 1 MELAALCRWGLLALLAPPGAASSTQVCTGTDMKRLRSPASPTHLDMLRLHYOGQVVOGNL 60
Db 1 MELAALCRWGLLALLAPPGAASSTQVCTGTDMKRLRSPASPTHLDMLRLHYOGQVVOGNL 60
Qy 61 ELTYLPTNASLSFLQDIQEVQGVLLIAHNOVRQVPLQRLRIVRGTQFDPEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIQEVQGVLLIAHNOVRQVPLQRLRIVRGTQFDPEDNYALAVLDNG 120
Qy 121 DPLNNTTPTVGTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
Db 121 DPLNNTTPTVGTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
Qy 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGGCARCKGPLPTDCCHQC 240
Db 181 LTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGGCARCKGPLPTDCCHQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPGRYTTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPGRYTTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR-----GTHSLLPRAAIVPVP 355
Db 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARCYGLGMEHLREVAVTSAN 360
Qy 356 LRMQPG--PAHPVLSFLRPSNDLVSAFYSLPLAPLSPTSVP-----SPVSVGRGPD 405
Db 361 IQBFAGCKKIFGSLAFPLSPFSGDPAASNT---APLQPEQLQVFTLEITGYLYISAWPD 417
Qy 406 --PDAHVAVNLSRYEG 419
Db 418 SLFDLSVFQNLQVIRG 433

RESULT 6
US-08-414-417B-68
; Sequence 68, Application US/08414417B
; Patent No. 5801095
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,417B
; FILING DATE: 31-MAR-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 622-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids

```
;
; TYPE: amino acid
; TOPOLOGY: linear
US-08-414-417B-68

Query Match
Best Local Similarity 82.1%; Score 1878; DB 1; Length 1255;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLALLPFGAASQVCTGTDKMLRLPASPETHLDMRLHYQGCQVQGNL 60
DB 1 MELAALCRWGLLALLPFGAASQVCTGTDKMLRLPASPETHLDMRLHYQGCQVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVLIHNRQVPLQRLIRVRGTQLPEDNVALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVLIHNRQVPLQRLIRVRGTQLPEDNVALAVLNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMKSGRCWGESSEDCOSLRTVTCAGGCARCKGPLETDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMKSGRCWGESSEDCOSLRTVTCAGGCARCKGPLETDCHEQC 240
QY 241 AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMNPPEGRTYFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMNPPEGRTYFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNRQVTAEDGTQRCCKSKPCAR-----GTHSLPRPAAVPPV 355
DB 301 YNYLSTDVGSCTLVCPHNRQVTAEDGTQRCCKSKPCAR-----GTHSLPRPAAVPPV 355
QY 356 LRMQPG--PAHPVLSFLRPSWDLVSFAFYLPLAPLSPTSVP1-----SPVSVGRGPD 405
DB 356 LRMQPG--PAHPVLSFLRPSWDLVSFAFYLPLAPLSPTSVP1-----SPVSVGRGPD 405
QY 361 IOEFAGCKKIFGSLAFPLESFDGDPASNT---APLQPEQLQVFTLEITGYLISAWPD 417
DB 361 IOEFAGCKKIFGSLAFPLESFDGDPASNT---APLQPEQLQVFTLEITGYLISAWPD 417
QY 406 --PDHVAVNLSRYEG 419
DB 418 SLPDLVSFQNLQVIRG 433

RESULT 8
US-08-486-348A-68
; Sequence 68, Application US/08486348A
; Patent No. 5846538
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disib, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69

Query Match
Best Local Similarity 82.1%; Score 1878; DB 2; Length 1255;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLALLPFGAASQVCTGTDKMLRLPASPETHLDMRLHYQGCQVQGNL 60
DB 1 MELAALCRWGLLALLPFGAASQVCTGTDKMLRLPASPETHLDMRLHYQGCQVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVLIHNRQVPLQRLIRVRGTQLPEDNVALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVLIHNRQVPLQRLIRVRGTQLPEDNVALAVLNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDNRSRACHPCSPMKSGRCWGESSEDCOSLRTVTCAGGCARCKGPLETDCHEQC 240
DB 181 LTLIDNRSRACHPCSPMKSGRCWGESSEDCOSLRTVTCAGGCARCKGPLETDCHEQC 240
QY 241 AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMNPPEGRTYFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMNPPEGRTYFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNRQVTAEDGTQRCCKSKPCAR-----GTHSLPRPAAVPPV 355
DB 301 YNYLSTDVGSCTLVCPHNRQVTAEDGTQRCCKSKPCAR-----GTHSLPRPAAVPPV 355
QY 356 LRMQPG--PAHPVLSFLRPSWDLVSFAFYLPLAPLSPTSVP1-----SPVSVGRGPD 405
DB 356 LRMQPG--PAHPVLSFLRPSWDLVSFAFYLPLAPLSPTSVP1-----SPVSVGRGPD 405
QY 361 IOEFAGCKKIFGSLAFPLESFDGDPASNT---APLQPEQLQVFTLEITGYLISAWPD 417
DB 361 IOEFAGCKKIFGSLAFPLESFDGDPASNT---APLQPEQLQVFTLEITGYLISAWPD 417
QY 406 --PDHVAVNLSRYEG 419
DB 418 SLPDLVSFQNLQVIRG 433

RESULT 8
US-08-486-348A-68
; Sequence 68, Application US/08486348A
; Patent No. 5846538
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disib, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
```


RESULT 10
US-08-468-545B-68
Sequence 68, Application US/08468545B
Patent No. 5876712
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,545B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C5
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-468-545B-68

Query Match 82.1%; Score 1878; DB 2; Length 1255;
Best Local Similarity 83.0%; Pred. No. 9.1e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLLALLPPGAASSTQVCTGDMKRLPASPETHLDMRLHLYQGCQVVGNL 60
DB 1 MELAALCRWGLLLALLPPGAASSTQVCTGDMKRLPASPETHLDMRLHLYQGCQVVGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVLIHAHQVQVPLQRLIRVGTQLFEDNYALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVLIHAHQVQVPLQRLIRVGTQLFEDNYALAVLNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
QY 181 LTLIDTNRSRACHPCSPCKGSRGWGSESSDCQSLTRTVCGAGGACRGKGLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPCKGSRGWGSESSDCQSLTRTVCGAGGACRGKGLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNLNQEVTAEDGTQRCCKSPCAR-----GTHSLPRPAVPVP 355
DB 301 YNYLSTDVGSCTLVCPHNLNQEVTAEDGTQRCCKSPCAR-----GTHSLPRPAVPVP 355
QY 356 LSWQPG--PAHPVLSFLRPSWDLVSAFYSLPLAPLSPTSVPI-----SPVSVGRGPD 405
DB 361 IQEFGACKKIFGSLAFLPSFGDPASNT---APLQPEQLQVFETLELTGLYISAFED 417

QY 406 --PDAHVAVNLSRYEG 419
DB 418 SLFDLSVFNQLQVIRG 433

RESULT 11
US-08-356-786-2
Sequence 2, Application US/08356786
Patent No. 5977305
GENERAL INFORMATION:
APPLICANT: Huston, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. L.
APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
TITLE OF INVENTION: Marker
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/831,967
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-786-2

Query Match 82.1%; Score 1878; DB 2; Length 1255;
Best Local Similarity 83.0%; Pred. No. 9.1e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLLALLPPGAASSTQVCTGDMKRLPASPETHLDMRLHLYQGCQVVGNL 60
DB 1 MELAALCRWGLLLALLPPGAASSTQVCTGDMKRLPASPETHLDMRLHLYQGCQVVGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVLIHAHQVQVPLQRLIRVGTQLFEDNYALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVLIHAHQVQVPLQRLIRVGTQLFEDNYALAVLNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
QY 181 LTLIDTNRSRACHPCSPCKGSRGWGSESSDCQSLTRTVCGAGGACRGKGLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPCKGSRGWGSESSDCQSLTRTVCGAGGACRGKGLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 355
Db 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 360
QY 356 LRMQPG--PAHPVLSFLRPSMDLVSFAFYSLPLAPLSPTSVP1-----SPVSVGRGPD 405
Db 361 IQEFAGCKKIFGSLAFLPESFDGDPASNT---APLQPEQLQVFTLEETITGYLYISAWPD 417
QY 406 --PDHVAVNLRYEG 419
Db 418 SLPDLVSFQNLQVIRG 433

RESULT 12
US-08-466-680B-68
; Sequence 68, Application US/08466680B
; Patent No. 6075122
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,680B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-466-680B-68

Query Match 82.1%; Score 1878; DB 3; Length 1255;
Best Local Similarity 83.0%; Pred. No. 9.1e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPETHLDMLRHLVGGCVVQGNL 60
Db 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPETHLDMLRHLVGGCVVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVLIHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 120
Db 61 ELTYLPTNASLSFLQDIQEVQGVLIHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 120
QY 121 DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
Db 121 DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPMCKSGRCWGSSESDCSLRTVTCAGGCARCKGPLPTDCCHQEC 240
Db 181 LTLIDTNRSRACHPCSPMCKSGRCWGSSESDCSLRTVTCAGGCARCKGPLPTDCCHQEC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 355
Db 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 360
QY 356 LRMQPG--PAHPVLSFLRPSMDLVSFAFYSLPLAPLSPTSVP1-----SPVSVGRGPD 405
Db 361 IQEFAGCKKIFGSLAFLPESFDGDPASNT---APLQPEQLQVFTLEETITGYLYISAWPD 417
QY 406 --PDHVAVNLRYEG 419

QY 181 LTLIDTNRSRACHPCSPMCKSGRCWGSSESDCSLRTVTCAGGCARCKGPLPTDCCHQEC 240
Db 181 LTLIDTNRSRACHPCSPMCKSGRCWGSSESDCSLRTVTCAGGCARCKGPLPTDCCHQEC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 355
Db 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 360
QY 356 LRMQPG--PAHPVLSFLRPSMDLVSFAFYSLPLAPLSPTSVP1-----SPVSVGRGPD 405
Db 361 IQEFAGCKKIFGSLAFLPESFDGDPASNT---APLQPEQLQVFTLEETITGYLYISAWPD 417
QY 406 --PDHVAVNLRYEG 419
Db 418 SLPDLVSFQNLQVIRG 433

RESULT 13
US-09-527-487-2
; Sequence 2, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-527-487-2

Query Match 82.1%; Score 1878; DB 4; Length 1255;
Best Local Similarity 83.0%; Pred. No. 9.1e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPETHLDMLRHLVGGCVVQGNL 60
Db 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPETHLDMLRHLVGGCVVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVLIHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 120
Db 61 ELTYLPTNASLSFLQDIQEVQGVLIHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 120
QY 121 DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
Db 121 DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPMCKSGRCWGSSESDCSLRTVTCAGGCARCKGPLPTDCCHQEC 240
Db 181 LTLIDTNRSRACHPCSPMCKSGRCWGSSESDCSLRTVTCAGGCARCKGPLPTDCCHQEC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 355
Db 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCVTACP 360
QY 356 LRMQPG--PAHPVLSFLRPSMDLVSFAFYSLPLAPLSPTSVP1-----SPVSVGRGPD 405
Db 361 IQEFAGCKKIFGSLAFLPESFDGDPASNT---APLQPEQLQVFTLEETITGYLYISAWPD 417
QY 406 --PDHVAVNLRYEG 419

Db 418 SLPDLVSFQNLQVIRG 433

RESULT 14
US-09-811-115-3
; Sequence 3, Application US/0981115
; Patent No. 6632979
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Erickson, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT 034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3

Query Match 82.1%; Score 1878; DB 4; Length 1255;
Best Local Similarity 83.0%; Pred. No. 9.1e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRGGLLALPPGAASQTCTGDMKLRIPASPTHLDMLRHLHYQCCVQGNL 60
DB 1 MELAALCRGGLLALPPGAASQTCTGDMKLRIPASPTHLDMLRHLHYQCCVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVLIHQNVRQVPLQRLRIVRGTLQFEDNYALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVLIHQNVRQVPLQRLRIVRGTLQFEDNYALAVLNG 120
QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
QY 181 LTLIDTNRGRACHPCSPMKGSRGCGESSEDQSLTRTVCCAGGCAKCGPLPTDCCHQC 240
DB 181 LTLIDTNRGRACHPCSPMKGSRGCGESSEDQSLTRTVCCAGGCAKCGPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSPCAR-----GTHSLPRPAAPVP 355
DB 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSPCAR-----GTHSLPRPAAPVP 355
QY 361 IQBFAGCKKIFGSLAFLPESFGDPSANT---APLQPEQLQVFTLEITGLYISAWPD 417
DB 361 IQBFAGCKKIFGSLAFLPESFGDPSANT---APLQPEQLQVFTLEITGLYISAWPD 417

RESULT 15
US-09-354-533-68
; Sequence 68, Application US/09354533
; Patent No. 6664370
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/354,533
; FILING DATE: 15-Jul-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-09-354-533-68

Query Match 82.1%; Score 1878; DB 4; Length 1255;
Best Local Similarity 83.0%; Pred. No. 9.1e-156;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRGGLLALPPGAASQTCTGDMKLRIPASPTHLDMLRHLHYQCCVQGNL 60
DB 1 MELAALCRGGLLALPPGAASQTCTGDMKLRIPASPTHLDMLRHLHYQCCVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVLIHQNVRQVPLQRLRIVRGTLQFEDNYALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVLIHQNVRQVPLQRLRIVRGTLQFEDNYALAVLNG 120
QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
QY 181 LTLIDTNRGRACHPCSPMKGSRGCGESSEDQSLTRTVCCAGGCAKCGPLPTDCCHQC 240
DB 181 LTLIDTNRGRACHPCSPMKGSRGCGESSEDQSLTRTVCCAGGCAKCGPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSPCAR-----GTHSLPRPAAPVP 355
DB 301 YNYLSTDVGSCTLVCPHLNQVTAEDGTQRCCKSPCAR-----GTHSLPRPAAPVP 355
QY 361 IQBFAGCKKIFGSLAFLPESFGDPSANT---APLQPEQLQVFTLEITGLYISAWPD 417
DB 361 IQBFAGCKKIFGSLAFLPESFGDPSANT---APLQPEQLQVFTLEITGLYISAWPD 417

RESULT 16
US-09-422-108-1
; Sequence 1, Application US/08422108
; Patent No. 6015667
; GENERAL INFORMATION:

APPLICANT: Hudziak, Robert M.
APPLICANT: Shepard, H. Michael
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,108
FILING DATE: 14-Apr-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/355460
FILING DATE: 13-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/048346
FILING DATE: 15-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/354319
FILING DATE: 19-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 554C2D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-422-108-1

Query Match 77.4%; Score 1769; DB 3; Length 624;
Best Local Similarity 82.2%; Pred. No. 1.3e-146;
Matches 341; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
QY 22 STQVCTGDMKRLPASPETHLDMLRHLYQGCVVQGNLELYLPTNASLSFLQDIQEVQ 81
DB 1 STQVCTGDMKRLPASPETHLDMLRHLYQGCVVQGNLELYLPTNASLSFLQDIQEVQ 60
QY 82 GYVLIAHNQVRQVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTPTVTGASPGGLRELQ 141
DB 61 GYVLIAHNQVRQVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTPTVTGASPGGLRELQ 120
QY 142 LRSLEILKGGVLIQNPOLCYQDTILWKDIFHKNQALTLIDNRSRACHPCSPMCKG 201
DB 121 LRSLEILKGGVLIQNPOLCYQDTILWKDIFHKNQALTLIDNRSRACHPCSPMCKG 180
QY 202 SRCWGESSEDCSLRTFTVCAGGCARCKPLPDDCCEQCAAGCTGPKHSDCLCLHFNHS 261
DB 181 SRCWGESSEDCSLRTFTVCAGGCARCKPLPDDCCEQCAAGCTGPKHSDCLCLHFNHS 240
QY 262 GICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLSLTDVGSCTLVCPHLNQE 321
DB 241 GICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLSLTDVGSCTLVCPHLNQE 300
QY 322 VTAEDGTQRCCKSKFCAR-----GTHSLPRAVPELRLNQCQ--PAHPVLSPILRSW 374
DB 301 VTAEDGTQRCCKSKFCARVCYGLGMEHLREVRVTSANIOFAGCKKIFGSLAFLPSPF 360

QY 375 DLVSIFYSLPLAPLSPTSVPI-----SPVSVGRGPD--PDHVAVNLSRYEG 419
DB 361 DGDPAISNT---APLOPEQLQVFFLEITEILYLYISAMPDLSPLDSVFONLQVIRG 412
RESULT 17
US-08-422-734-1
; Sequence 1, Application US/08422734
; Patent No. 6333169
; GENERAL INFORMATION:
; APPLICANT: Hudziak, Robert M.
; APPLICANT: Shepard, H. Michael
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,734
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/422108
; FILING DATE: 14-Apr-1995
; APPLICATION NUMBER: 08/355460
; FILING DATE: 13-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/048346
; FILING DATE: 15-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/354319
; FILING DATE: 19-MAY-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 554C2D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 624 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-422-734-1

Query Match 77.4%; Score 1769; DB 4; Length 624;
Best Local Similarity 82.2%; Pred. No. 1.3e-146;
Matches 341; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
QY 22 STQVCTGDMKRLPASPETHLDMLRHLYQGCVVQGNLELYLPTNASLSFLQDIQEVQ 81
DB 1 STQVCTGDMKRLPASPETHLDMLRHLYQGCVVQGNLELYLPTNASLSFLQDIQEVQ 60
QY 82 GYVLIAHNQVRQVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTPTVTGASPGGLRELQ 141
DB 61 GYVLIAHNQVRQVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTPTVTGASPGGLRELQ 120
QY 142 LRSLEILKGGVLIQNPOLCYQDTILWKDIFHKNQALTLIDNRSRACHPCSPMCKG 201
DB 121 LRSLEILKGGVLIQNPOLCYQDTILWKDIFHKNQALTLIDNRSRACHPCSPMCKG 180

QY 202 SRCWGESSEDQSLTRTVCAAGGCKGPIPTDCCHEQCAAGCTGPKHSDCLACLFHNS 261
DB 181 SRCWGESSEDQSLTRTVCAAGGCKGPIPTDCCHEQCAAGCTGPKHSDCLACLFHNS 240
QY 262 GICELHCPALVTNTDTFESMPNPEGRYTFGASCVTACPNYLSTDVGSCTIVCPLEHNS 321
DB 241 GICELHCPALVTNTDTFESMPNPEGRYTFGASCVTACPNYLSTDVGSCTIVCPLEHNS 300
QY 322 VTAEDGTQCEKSKPCAR-----GTHSLLPRAAIVPVLKMPG--PAHVLSPFLRPSW 374
DB 301 VTAEDGTQCEKSKPCARVYGLGMEHLREAVTSAIQEPAGCKKIFGSLAFLPESF 360
QY 375 DLVSAPVSLPLAFLSPTSVPI-----SPVSVGRGPD--PDHVAVNLRYEG 419
DB 361 DGDPAANT---APLQPEQLQVFTLEETIYLYISAWPSLPLDSVFNQLQVIRG 412

RESULT 18

US-09-648-067A-1
; Sequence 1, Application US/09648067A
; Patent No. 6627196
; GENERAL INFORMATION:
; APPLICANT: Baughman, Sharon A.
; TITLE OF INVENTION: Dosages for Treatment with Anti-ErbB2 Antibodies
; FILE REFERENCE: P1775R1
; CURRENT APPLICATION NUMBER: US/09/648,067A
; PRIOR FILING DATE: 2000-08-25
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: US 60/213,822
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 1
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-648-067A-1

Query Match 37.6%; Score 860; DB 4; Length 166;
Best Local Similarity 100.0%; Pred. No. 9.5e-68;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26 CTGTDMLRLPASPETHLMLRLHYQGVVQGNLELYLPTNASLSFLQDIQEVQGYVL 85
DB 1 CTGTDMLRLPASPETHLMLRLHYQGVVQGNLELYLPTNASLSFLQDIQEVQGYVL 60
QY 86 IAHNVQVPLQRLIRVGTQLPEDNYALAVLDNGDPLANTTPVTGASPGGLRLQLRSL 145
DB 61 IAHNVQVPLQRLIRVGTQLPEDNYALAVLDNGDPLANTTPVTGASPGGLRLQLRSL 120
QY 146 TELKGGVLIQNPOLCYQDTILWKDIFHKNNQALTLIDTNRSA 191
DB 121 TELKGGVLIQNPOLCYQDTILWKDIFHKNNQALTLIDTNRSA 166

RESULT 19

US-08-336-708A-9
; Sequence 9, Application US/08336708A
; Patent No. 5521295
; GENERAL INFORMATION:
; APPLICANT: Pacifici, Robert E.
; APPLICANT: Thomson, Arlen R.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: Hybrid Receptor Molecules
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,708A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oleski, Nancy
REFERENCE/DOCKET NUMBER: A-241A
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 644 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-708A-9
Query Match 34.7%; Score 793; DB 1; Length 644;
Best Local Similarity 45.3%; Pred. No. 4.1e-61; Indels 14; Gaps 4;
Matches 151; Conservative 50; Mismatches 118;
QY 11 LLLALLPPGAA--STQVCTGTDMLRLPASPETHLMLRLHYQGVVQGNLELYLPTN 68
DB 14 LLAALCPASRALEEKVKCGTISNKLTLQGTEDHFLSLQRMFNCEVVLGNLELYVQVN 73
QY 69 ASLSFLQDIQEVQGYVLIAHNVQVPLQRLIRVGTQLPEDNYALAVLDNGDPLANTTP 128
DB 74 YDLSFLKTIQEVAGVYLIATNTVERIPLENLQINGNMYSYALAVLSNYD----- 126
QY 129 VTGASPGGLRLQLRSLIRLTKLGGVLIQNPOLCYQDTILWKDIFHKNNQALTLIDTNR 188
DB 127 ---ANKTGLKELPMENLQELLEGAVRFSNNPALCNVESIQWRDIVSSDFLSNMSDFQNH 183
QY 189 SRACHPCSPMKGSRCSWGESSEDCOSLRTRTVCAAGCA-RCKGPLPTDCCHEQCAAGCTGP 247
DB 184 LGSCQKCDPSPNGSCWGAEGENCQKLTIIAQQSCGRGCRKSPSCCHNCAAGCTGP 243
QY 248 KHSQCLAGLHNHSGICELHCPALVTNTDTFESMPNPEGRYTFGASCVTACPNYLST 307
DB 244 RESDCLVCRFRDEATCKOTCPPLMLYNTTYQMDVNEFGKYSFGATCVKCKPRYVVD 303
QY 308 VGSCTIVCPLEHNSQVTAEDGTQCEKSKPCAR 340
DB 304 HGSCVRAACADSYEM-ESDGVKCKKCKGCPCK 335

RESULT 20

US-08-484-438-7
; Sequence 7, Application US/08484438
; Patent No. 5811098
; Patent No. 5811098 5780031
; GENERAL INFORMATION:
; APPLICANT: Plowman, Gregory D.
; APPLICANT: Culouscou, Jean-Michel
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Siegall, Clay B.
; APPLICANT: Hellstr m, Ingegerd
; APPLICANT: Hellstr m, Karl E.
; TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,438
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,442
FILING DATE: 14-OCT-1994
APPLICATION NUMBER: US 08/150,704
FILING DATE: 10-NOV-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/981,165
FILING DATE: 24-NOV-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9030
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1210 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-438-7

Query Match 34.7%; Score 793; DB 2; Length 1210;
Best Local Similarity 45.3%; Pred. No. 9.5e-61;
Matches 151; Conservative 50; Mismatches 118; Indels 14; Gaps 4;

QY 11 LLLALPPGAA--STQVCTGTDMLRLPASPTHLDMLRLHYQGVQVQGNLEITYPTN 68
DB 14 LLAALCPASRALEKKVCGQTSNKLTLQGTGFEDHLSLQRMFNCEVVLGNLEITYVQRN 73
QY 69 ASLSFLQDIQEVGVVLIANQVPLQRLIRVGTQFQEDNVALAVLDNGDPLNNTTP 128
DB 74 YDLSFLKTIQEVAGVYLIANTVERIPLENQIIRGNMYNYSVALAVLSYD----- 126
QY 129 VTGASPGGLRELQRLSLEILKGGVLIQRPOLCYQDTILWKDIFHKNQALTLIDTNR 188
DB 127 ---ANKTLKELPMRLQELIHGAVRFSNNPFCNVESIQWRDIVSSDFLSNMMDPQNH 183
QY 189 SRACHPCSPMCKGSRGWESSEDCQSLTRTVACAGCA-RCKGPLPTDCHQCAAGCTGP 247
DB 184 LGSCQKCDPSPGSCGAGGECNCKLTKIICAOQCSGRCKGKSPSCCHNQCAAGCTGP 243
QY 248 KHSDDLACLHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYVLTSD 307
DB 244 RESDCLVCRKFRDEATCKDTCPLMLYNTTYQMDVNPCKYSGFATCVKCKCPRYVYTD 303
QY 308 VGSCTLVCLHNEQVTAEDGTQRCCKSKPCAR 340
DB 304 HGSCVRACGADSYEM-EEDGVRKCKCKGCPCK 335

RESULT 22
US-08-484-438-10
Sequence 10, Application US/08484438
Patent No. 5811098
Patent No. 5811098 5780031
GENERAL INFORMATION:
APPLICANT: Plowman, Gregory D.
APPLICANT: Culouscou, Jean-Michel
APPLICANT: Shoyab, Mohamed
APPLICANT: Siegall, Clay B.
APPLICANT: Hellstr m, Ingegerd
APPLICANT: Hellstr m, Karl E.

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,438
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,442
FILING DATE: 14-OCT-1994
APPLICATION NUMBER: US 08/150,704
FILING DATE: 10-NOV-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/981,165
FILING DATE: 24-NOV-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9030
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1210 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-438-7

Query Match 34.7%; Score 793; DB 2; Length 1210;
Best Local Similarity 45.3%; Pred. No. 9.5e-61;
Matches 151; Conservative 50; Mismatches 118; Indels 14; Gaps 4;

QY 11 LLLALPPGAA--STQVCTGTDMLRLPASPTHLDMLRLHYQGVQVQGNLEITYPTN 68
DB 14 LLAALCPASRALEKKVCGQTSNKLTLQGTGFEDHLSLQRMFNCEVVLGNLEITYVQRN 73
QY 69 ASLSFLQDIQEVGVVLIANQVPLQRLIRVGTQFQEDNVALAVLDNGDPLNNTTP 128
DB 74 YDLSFLKTIQEVAGVYLIANTVERIPLENQIIRGNMYNYSVALAVLSYD----- 126
QY 129 VTGASPGGLRELQRLSLEILKGGVLIQRPOLCYQDTILWKDIFHKNQALTLIDTNR 188
DB 127 ---ANKTLKELPMRLQELIHGAVRFSNNPFCNVESIQWRDIVSSDFLSNMMDPQNH 183
QY 189 SRACHPCSPMCKGSRGWESSEDCQSLTRTVACAGCA-RCKGPLPTDCHQCAAGCTGP 247
DB 184 LGSCQKCDPSPGSCGAGGECNCKLTKIICAOQCSGRCKGKSPSCCHNQCAAGCTGP 243
QY 248 KHSDDLACLHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYVLTSD 307
DB 244 RESDCLVCRKFRDEATCKDTCPLMLYNTTYQMDVNPCKYSGFATCVKCKCPRYVYTD 303
QY 308 VGSCTLVCLHNEQVTAEDGTQRCCKSKPCAR 340
DB 304 HGSCVRACGADSYEM-EEDGVRKCKCKGCPCK 335

RESULT 21
US-08-475-035-4
Sequence 4, Application US/08475035
Patent No. 5985553
GENERAL INFORMATION:
APPLICANT: KING, C. R.
APPLICANT: KRAUS, MATTHIAS H.
APPLICANT: AARONSON, STUART A.
TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM
TITLE OF INVENTION: EGF RECEPTOR GENE

TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE

NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,438
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/323,442
 FILING DATE: 14-OCT-1994
 APPLICATION NUMBER: US 08/150,704
 FILING DATE: 10-NOV-1993
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/981,165
 FILING DATE: 24-NOV-1992
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Mirock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 5624-230
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 10:
 LENGTH: 911 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-484-438-10

Query Match 33.9%; Score 775; DB 2; Length 911;
 Best Local Similarity 45.1%; Pred. No. 2.5e-59;
 Matches 157; Conservative 45; Mismatches 124; Indels 22; Gaps 7;

QY 9 WGLLLALLPPGAA---STQVCTGTDMLRLPASPEHLDMLRHLYGQGVQVGNLELTY 64
 DB 8 WVVSLVLAAGTVQPSQSVCACTENKLSLSLEQQYRALRYKYENCVWGNLEITS 67

QY 65 LPTNASLSFLQDIQEVGYVLIHQNQVRQVPLQRLRIVRGTLQFEDNYALAVLDNGDPLN 124
 DB 68 IEHRDLSFLRSVREVTGYVLVALNQFVLPLENLRIINGTKLYEDRYALAIFLNYRKDG 127

QY 125 NTTPVTGASPGGLRLQLRSITLTKGVLIQNPOLCYODTILWKDIFHKNQALTLI 184
 DB 128 NF-----GLOELGKLNLTILNGGVYVDQNKFLCYADTIHQDIVRNWPSNLTIV 178

QY 185 DITNRSRACHPCSPMKGRSGESEDQSLTRTVAGGC-ARCKGPLPTDCHEQCAAG 243
 DB 179 STNGSSGGRCHKSCG-RCMGPTENHCQTLTRTVCAEQDGRGCVGYVSDCHRECAGG 237

QY 244 CTGPKHSDCLAHFNHSGICELHCPALVTNTDTESMPNPEGRYTFGASCVTACPNY 303
 DB 238 CSGPOTDTCFACMNFDSGACVTCQPTQFVFNPTTFQLEHFNKRYTYGAFVKKCPHF 297

QY 304 LSTDVGSCTILVCPILHNOEVTAEQDTCRCKSKPCAR-----GTHSL 346
 DB 298 V-VDSGSCVRCAPSSKMEV-EENGKCKPCCTDICKACDGIQGTGSLM 343

RESULT 23

US-08-484-438-4
 Sequence 4, Application US/08484438
 Patent No. 5811098
 Patent No. 5811098 5780031
 GENERAL INFORMATION:
 APPLICANT: Plowman, Gregory D.
 APPLICANT: Culouscou, Jean-Michel
 APPLICANT: Shoyab, Mohammed
 APPLICANT: Siegall, Clay B.
 APPLICANT: Hallstr m, Inggerd
 APPLICANT: Hellstr m, Karl E.
 TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,438
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/323,442
 FILING DATE: 14-OCT-1994
 APPLICATION NUMBER: US 08/150,704
 FILING DATE: 10-NOV-1993
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/981,165
 FILING DATE: 24-NOV-1992
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Mirock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 5624-230
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1058 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-484-438-4

Query Match 33.9%; Score 775; DB 2; Length 1058;
 Best Local Similarity 45.1%; Pred. No. 3e-59;
 Matches 157; Conservative 45; Mismatches 124; Indels 22; Gaps 7;

QY 9 WGLLLALLPPGAA---STQVCTGTDMLRLPASPEHLDMLRHLYGQGVQVGNLELTY 64
 DB 8 WVVSLVLAAGTVQPSQSVCACTENKLSLSLEQQYRALRYKYENCVWGNLEITS 67

QY 65 LPTNASLSFLQDIQEVGYVLIHQNQVRQVPLQRLRIVRGTLQFEDNYALAVLDNGDPLN 124
 DB 68 IEHRDLSFLRSVREVTGYVLVALNQFVLPLENLRIINGTKLYEDRYALAIFLNYRKDG 127

QY 125 NTTPVTGASPGGLRLQLRSITLTKGVLIQNPOLCYODTILWKDIFHKNQALTLI 184
 DB 128 NF-----GLOELGKLNLTILNGGVYVDQNKFLCYADTIHQDIVRNWPSNLTIV 178

QY 185 DTRSRACHPCSPCKGSRGWGSESDCQSLTRTVCAAGC-ARCKGFLPTDCHCQCAAG 243
DB 179 STNGSGGCRCHKSCGTG-RWGPTEHHCQTLTRVCAEQCDRCYCPYVSDCHRECAAG 237
QY 244 CTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPVNY 303
DB 238 CSFPGKDTDCFAWPNFNSGACVTCQCPOTFYNTTFQLEHFNFAKTYTGAFCKVCKCPNF 297
QY 304 LSTDVGSCTLVCPHNSQVTAEDGTORCEKCKPCAR-----GTHSL 346
DB 298 V-VDSSSCVRACPSKMEV-EENGIRKMKCPCTDICKPCACDGIOTGSLM 343

RESULT 24

US-08-484-438-2
; Sequence 2, Application US/08484438
; Patent No. 5811098
; Patent No. 5811098 5780031
; GENERAL INFORMATION:
; APPLICANT: Plowman, Gregory D.
; APPLICANT: Culouscou, Jean-Michel
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Siegall, Clay B.
; APPLICANT: Hellstz m, Ingegerd
; APPLICANT: Hellstz m, Karl B.
; TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,438
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/323,442
; FILING DATE: 14-OCT-1994
; APPLICATION NUMBER: US 08/150,704
; FILING DATE: 10-NOV-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/981,165
; FILING DATE: 24-NOV-1992
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,972
; REFERENCE/DOCKET NUMBER: 5624-230
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1308 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-484-438-2

Query Match 33.9%; Score 775; DB 2; Length 1308;
Best Local Similarity 45.1%; P-red. No. 4e-59;
Matches 157; Conservative 45; Mismatches 124; Indels 22; Gaps 7;

QY 9 WGLLALLPPGAA-----STQVCTGTDMLRLPASPTHLDMLRHLVQGVQVQGNLELTY 64
DB 8 WWSVSLVAAAGTVQPSDSQSCAGTENKLSLSLEQOYALARKYVENEVWGNLEITS 67
QY 65 LPTVASLFDLODQEVQGVYLIHNNQVQVPLQRLRIVRGTLQFEDNYALAVLNDGSDPLN 124
DB 68 IEHRDLSFURSREVTGYVLVALNQFRYLPENLRIRGTLKYEDRYALAIPIFYRKDG 127
QY 125 NTPPVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTTILWKDIFHKNNQLALTIL 184
DB 128 NF-----GLQELGLKMLTEILNGVVYVDQNKFLCYADTIHWQDIVRNPNPNSLTLV 178
QY 185 DTRSRACHPCSPCKGSRGWGSESDCQSLTRTVCAAGC-ARCKGFLPTDCHCQCAAG 243
DB 179 STNGSGGCRCHKSCGTG-RWGPTEHHCQTLTRVCAEQCDRCYCPYVSDCHRECAAG 237
QY 244 CTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPVNY 303
DB 238 CSFPGKDTDCFAWPNFNSGACVTCQCPOTFYNTTFQLEHFNFAKTYTGAFCKVCKCPNF 297
QY 304 LSTDVGSCTLVCPHNSQVTAEDGTORCEKCKPCAR-----GTHSL 346
DB 298 V-VDSSSCVRACPSKMEV-EENGIRKMKCPCTDICKPCACDGIOTGSLM 343

RESULT 25

US-09-570-454-2
; Sequence 2, Application US/09570454
; Patent No. 6393743
; GENERAL INFORMATION:
; APPLICANT: Department of Veterans Affairs
; TITLE OF INVENTION: Isolation and characterization of epidermal growth
; FILE OF INVENTION: factor related protein
; FILE REFERENCE: 107999.00106
; CURRENT APPLICATION NUMBER: US/09/570,454
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: 60/134,200
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-570-454-2

Query Match 33.8%; Score 773; DB 4; Length 478;
Best Local Similarity 45.7%; P-red. No. 1.6e-59;
Matches 154; Conservative 43; Mismatches 118; Indels 22; Gaps 6;

QY 11 LLLALLPPGAA--STQVCTGTDMLRLPASPTHLDMLRHLVQGVQVQGNLELTYLPTN 68
DB 14 LUTALCAAGGALREKKVCGTSNRLTQJGTFEDHFLSLQRMVNNCEVLGNLEITYVQRN 73
QY 69 ASLSFLDLODQEVQGVYLIHNNQVQVPLQRLRIVRGTLQFEDNYALAVLNDGSDPLNNTTP 128
DB 74 YLSEFLTKIQEVAGYFLIALNTVERIPSEDQITINGNALYENTYALATLSN----- 124
QY 129 VTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTTILWKDI----FHKNQLALTIL 184
DB 125 -YGTNRTGLRELPMENLOELLIGAVRFSNPNILCNMDTIQWRDIVQNVFMSMDL--- 180
QY 185 DTRSRACHPCSPCKGSRGWGSESDCQSLTRTVCAAGC-ARCKGFLPTDCHCQCAAG 243
DB 181 -QSHFSSCPKCDPFCPNSGCGGEGENCKLTKIQAQCCSHRCRCRGRSPSCCHNQCAAG 239
QY 244 CTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPVNY 303
DB 240 CTGPKQKSDCLVCQKQFQDEATCKDTCFPLMLYNPTTYQVDAVPEKGYSGATCVKNCPRNY 299
QY 304 LSTDVGSCTLVCPHNSQVTAEDGTORCEKCKPCAR 340
DB 300 VVTDHSSCVRACGPDYIEV-BEDGIRKCKCKCDGCPCK 335


```
RESULT 26
US-09-867-521-2
; Sequence 2, Application US/09867521
; Patent No. 6582934
; GENERAL INFORMATION:
; APPLICANT: Department of Veterans Affairs
; TITLE OF INVENTION: Isolation and characterization of epidermal growth
; factor related protein
; FILE REFERENCE: 111828-00103
; CURRENT APPLICATION NUMBER: US/09/867,521
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/134,200
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/570,454
; PRIOR FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-867-521-2

Query Match      33.8%; Score 773; DB 4; Length 478;
Best Local Similarity 45.7%; Pred. No. 1.6e-59;
Matches 154; Conservative 43; Mismatches 118; Indels 22; Gaps 6;

QY 11 LLLALLPPGAA--STQVCTGTDMLRLPASPETHLMLRLHLYQCCVQGNLELTLYLPTN 68
DB 14 LTLALCAAGALEEKVCGGTSNRLTQGFDFHLSLQRMVNNCEVWIGNLEITYVORN 73
QY 69 ASLSFQDIQEVQGYVLIHNOVRQVPLQRLIRVGTQFEDNYALAVLNGDPLNNTTP 128
DB 74 YDLSFLKTIQEVAGYFLIALNTVERIPSEDQIIRGNALYENTYALAILSN----- 124
QY 129 VTGASPGGLRELQRLSRLTEILKGVLIQRNPQLCYQDTILWKDI-----FHNKQALTLI 184
DB 125 -YGTNRTGRELPMNLQILIGAVRPNPILCNNDTIQWRDVGQVFNMSNMDL---- 180
QY 185 DYNRSRACHPCSPKSCGSGWSESSDQSLRTVVCAGCA-RCKGPLPTDCHEQCAAG 243
DB 181 -QSHPSCKPCDFPCFNGCWGGGENCQKLIICAQCSHRGRSFSDCCHVQCAAG 239
QY 244 CTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCVTACPNY 303
DB 240 CTGPQKSDCLVCKQFQDEATCKOTCPPLMLYNPTTYQMDVNPBGKYSFGATCVGNCP 299
QY 304 LSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR 340
DB 300 VTDHSGVRACPDYIEV-EEDGIRKCKKCGPCR 335

RESULT 27
US-07-978-895-4
; Sequence 4, Application US/07978895
; Patent No. 5480968
; GENERAL INFORMATION:
; APPLICANT: Kraus, Matthias H.
; APPLICANT: Aazonson, Stuart A.
; TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
; EPIDERMAL GROWTH FACTOR RECEPTOR, ANTIGEN THEREOF, AND
; BIOASSAYS AND METHODS RELATED THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Suite 400
; STREET: 133 Carnegie Way, N.W.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.A.
; ZIP: 30303
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/978,895
; FILING DATE: 19921110
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/444,406
; FILING DATE: 01-DEC-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Perryman, David G.
; REGISTRATION NUMBER: 33,439
; REFERENCE/DOCKET NUMBER: 1414-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 688-0770
; TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-978-895-4

Query Match      32.2%; Score 735.5; DB 1; Length 1342;
Best Local Similarity 44.0%; Pred. No. 1.2e-55;
Matches 150; Conservative 49; Mismatches 121; Indels 21; Gaps 9;

QY 10 GLLIALLPPGAA--STQVCTGTDMLRLPASPETHLMLRLHLYQCCVQGNLELTLYLPT 67
DB 11 GLLFSLARGSEVGNQAVCPGLTGLSVTDAENQYQTLKLYRCEVVMGNLEIVLTGH 70
QY 68 NASLSFLQDIQEVQGYVLIHNOVRQVPLQRLIRVGTQFEDNYALAVLNGDPLNNTT 127
DB 71 NADLSFLQDIQEVQGYVLIHNOVRQVPLQRLIRVGTQFEDNYALAVLNGDPLNNTT 125
QY 128 PVTGASPGGLRELQRLSRLTEILKGVLIQRNPQLCYQDTILWKDI FHNKQALTLIDTN 187
DB 126 ---NSSHALRQLRLTQLTEILSGVYIEKNDLCHMDTDWRDVRDRD---ABIVKND 178
QY 188 RSRACHPCSPKSCGSGWSESSDQSLRTVVCAGGC-ARCKGPLPTDCHEQCAAGCTG 246
DB 179 NGRSPPCHEVCKG-RCWGPGSEDCQTLTKTICAPQCNHCFGFPNQCCHDECAGGCG 237
QY 247 PKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRTYFGASCVTACPNY 306
DB 238 PQTDFACRHFNDGACVPRCPQPLVYNKLTFOLEPNHTKYQYGGVGVASCPHNFV-V 296
QY 307 DVGSCCTLVCPHLNQEVTAEDGTQRCCKSKPCAR---GTHS 344
DB 297 DQTSVRACPPDKMEVD-KNGLKMWCEPGGLCPKACBGTGS 336

RESULT 28
US-08-484-438-9
; Sequence 9, Application US/08484438
; Patent No. 5811098
; Patent No. 5811098 5780031
; GENERAL INFORMATION:
; APPLICANT: Flowman, Gregory D.
; APPLICANT: Culouscou, Jean-Michel
; APPLICANT: Shovab, Mohammed
; APPLICANT: Siegall, Clay B.
; APPLICANT: Hellstr m, Ingegerd
; APPLICANT: Hellstr m, Karl E.
; TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
```

STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,438
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,442
FILING DATE: 14-OCT-1994
APPLICATION NUMBER: US 08/150,704
FILING DATE: 10-NOV-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/981,165
FILING DATE: 24-NOV-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1342 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-438-9

Query Match 32.2%; Score 735.5; DB 2; Length 1342;
Best Local Similarity 44.0%; Pred. No. 1.2e-55;
Matches 150; Conservative 49; Mismatches 121; Indels 21; Gaps 9;
Qy 10 GLLALLPPGAA--STQVCTGTDMKRLRLPASPTHLDMLRLHYGCGVQVGNLELYLPT 67
Db 11 GLFLSARGSEVGNQAVCPGTLNGLSVTGAENQYQTLKLYERCEVVMGNLELYLVTGH 70
Qy 63 NASLSFLQDIQRYQGVYLIANNOVQVLPQRLRIVRGTLQFEDNYALAVLDNGDPLNNTT 127
Db 71 NADLSFLQWIREVTGYVLVAMNEFSTLPLNLRVVRGTQVYDGKFAIFVM-----LNNT 125
Qy 128 PVTGASPGGLRELQRLSLTEILKGGVLIQPNPQLCYQDTILKWDIFHKNNQLALTLDTN 187
Db 126 ----NSSHALRQLRLTQLTEILSGVYIEKNDKLCHEMDTIDMRDIVRDRD---AEIVVKD 178
Qy 188 RSRACHPCSPCKSGRCWGESSEDCQSILTRTVGAGCC-ARCKGPLPTDCHEQCAAGCTG 246
Db 179 NGRSCPPCHEVCKG-RCWGPGSEDCQTLTKTICAPQCNHGCFGNPNQCCHECAGCGSG 237
Qy 247 PKHSDCLACHFNHSGICELHCPALVTYNTDIFESMNPENRGYTFGASCVTACPNYVLT 306
Db 238 PQDTDCFACRHFNDGACVPCPQPLVYKLTFLQLEPNPHTKYQYGGVGVASCFHNFV-V 296
Qy 307 DVGSCTLVCPHNEVTAEDGTQRCCKSKPCAR---GTHS 344
Db 297 DQTSVCVRACPPDKMEVD-KNGLKMCPECGGLCPKACEGTGS 336

RESULT 29

US-08-473-119-4

; Sequence 4, Application US/08473119
; Patent No. 5820859

; GENERAL INFORMATION:

APPLICANT: Kraus, Matthias H.
APPLICANT: Aaronson, Stuart A.
TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR RECEPTOR, ANTIGEN THEREOF, AND
TITLE OF INVENTION: BIOASSAYS AND METHODS RELATED THEREO
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Suite 400
STREET: 133 Carnegie Way, N.W.
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.A.
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,119
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/978,895
FILING DATE: 10-NOV-1992
APPLICATION NUMBER: US 07/444,406
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Perryman, David G.
REGISTRATION NUMBER: 33,438
REFERENCE/DOCKET NUMBER: 1414-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 688-0770
TELEFAX: (404) 688-9880
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1342 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-473-119-4

Query Match 32.2%; Score 735.5; DB 2; Length 1342;
Best Local Similarity 44.0%; Pred. No. 1.2e-55;
Matches 150; Conservative 49; Mismatches 121; Indels 21; Gaps 9;
Qy 10 GLLALLPPGAA--STQVCTGTDMKRLRLPASPTHLDMLRLHYGCGVQVGNLELYLPT 67
Db 11 GLFLSARGSEVGNQAVCPGTLNGLSVTGAENQYQTLKLYERCEVVMGNLELYLVTGH 70
Qy 68 NASLSFLQDIQRYQGVYLIANNOVQVLPQRLRIVRGTLQFEDNYALAVLDNGDPLNNTT 127
Db 71 NADLSFLQWIREVTGYVLVAMNEFSTLPLNLRVVRGTQVYDGKFAIFVM-----LNNT 125
Qy 128 PVTGASPGGLRELQRLSLTEILKGGVLIQPNPQLCYQDTILKWDIFHKNNQLALTLDTN 187
Db 126 ----NSSHALRQLRLTQLTEILSGVYIEKNDKLCHEMDTIDMRDIVRDRD---AEIVVKD 178
Qy 188 RSRACHPCSPCKSGRCWGESSEDCQSILTRTVGAGCC-ARCKGPLPTDCHEQCAAGCTG 246
Db 179 NGRSCPPCHEVCKG-RCWGPGSEDCQTLTKTICAPQCNHGCFGNPNQCCHECAGCGSG 237
Qy 247 PKHSDCLACHFNHSGICELHCPALVTYNTDIFESMNPENRGYTFGASCVTACPNYVLT 306
Db 238 PQDTDCFACRHFNDGACVPCPQPLVYKLTFLQLEPNPHTKYQYGGVGVASCFHNFV-V 296
Qy 307 DVGSCTLVCPHNEVTAEDGTQRCCKSKPCAR---GTHS 344
Db 297 DQTSVCVRACPPDKMEVD-KNGLKMCPECGGLCPKACEGTGS 336

RESULT 30

US-08-475-352-4

Search completed: July 4, 2004, 04:22:15
Job time : 29.7651 secs

Sequence 4, Application US/08475352
Patent No. 5916755
GENERAL INFORMATION:
APPLICANT: Kraus, Matthias H.
APPLICANT: Aaronson, Stuart A.
TITLE OF INVENTION: AN ISOLATED POLYPEPTIDE RELATED TO THE
TITLE OF INVENTION: EPIDERMAL GROWTH FACTOR RECEPTOR, ANTIGEN THEREOF, AND
TITLE OF INVENTION: BIOASSAYS AND METHODS RELATED THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Suite 400
STREET: 133 Carnegie Way, N.W.
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.A.
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,352
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/978,895
FILING DATE:
APPLICATION NUMBER: US 07/444,406
FILING DATE: 01-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Pettyman, David G.
REGISTRATION NUMBER: 33,438
REFERENCE/DOCKET NUMBER: 1414-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 688-0770
TELEFAX: (404) 688-9880
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1342 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-475-352-4

Query Match 32.2%; Score 735.5; DB 2; Length 1342;
Best Local Similarity 44.0%; Pred. No. 1.2e-55;
Matches 150; Conservative 49; Mismatches 121; Indels 21; Gaps 9;

Qy	10	GLLALLPGAA--STQVCVTGTDWKLRLFPSPETHLMLRHLHYGQGVQGNLELYLPT	67
Db	11	GLLFLSLARGSEVNGNSQAVCGTGLNGLSVTGDENQYQTLVLYRCEVWVGNLEIVLTGH	70
Qy	68	NASLSFLQDIQRYGVYLAHNOVQVPLQRLIRIVRGTLQFEDNYALAVLDNGDPLNNTT	127
Db	71	NADLSFLQIREVTGVYLVAMNEFTLPLPLNRVVRGTQYDQKFAIFVN-----LNVT	125
Qy	128	PVTGASPGGLREILQLRSLTEILKGGVLIQENPOLCYQDTILWKDIIPHKNQLALTLIDTN	187
Db	126	----NSSHALRQLRLQLTEILLGGVYIEKNDKLCHMDTIDWRDIVRDRD---AEIVWD	178
Qy	188	RSPACHPCSPMKSGSCWESSDDCSLRTVTCAGGC-ARCKGPIPTDCHQCQAGCTG	246
Db	179	NGRSCFPCHEVCCKG-KCWPGSGDCQTLTKTICAPQNGHCFGPNPQCCHDECAGGCSG	237
Qy	247	PKHSDCLACLFHFNHSGICSLHCFALVTYNTDTPESMPNPEGRYTFGASCVTACPNYLST	306
Db	238	PQDTDCFACEHFNDSACVPRCPQPLVYKLLTFQLEPNHTKYQYGGVCVASCPIHFV-V	296
Qy	307	DVGSCTLVPLHNQEVTAEDGTQRCCKSKPCAR---GTHS	344
Db	297	DQTSVCRACFPDRMEVD-KNGLKVCPCGGLCPKACEGTGS	336

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: July 4, 2004, 04:21:44 ; Search time 84.9779 Seconds
(without alignments)
1534.845 Million cell updates/sec

Title: US-09-506-079H-12
Perfect score: 2287
Sequence: 1 MELALCRWGLLIALLPPGA.....VGRGPDPAHVAVNLRSYEG 419
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 31283816 residues
Total number of hits satisfying chosen parameters: 1276540
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 90 summaries

- Database : Published Applications AA.*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCR_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description	
Result No.	Score	Query Match	ID	Length	DB	
1	2183	95.5	419	12	US-10-344-470-2	Sequence 2, Appli
2	2171	94.9	419	16	US-10-302-663-2	Sequence 2, Appli
3	1879.5	82.2	720	15	US-10-412-804A-6	Sequence 6, Appli
4	1878	82.1	645	9	US-09-921-161-1	Sequence 1, Appli
5	1878	82.1	645	14	US-10-268-501-13	Sequence 13, Appli
6	1878	82.1	645	15	US-10-608-626-13	Sequence 13, Appli
7	1878	82.1	653	9	US-09-854-356-3	Sequence 3, Appli
8	1878	82.1	685	15	US-10-412-804A-4	Sequence 4, Appli
9	1878	82.1	690	15	US-10-412-804A-11	Sequence 11, Appli
10	1878	82.1	712	9	US-09-854-356-7	Sequence 7, Appli
11	1878	82.1	715	15	US-10-412-804A-10	Sequence 10, Appli
12	1878	82.1	919	9	US-09-854-356-6	Sequence 6, Appli
13	1878	82.1	1253	14	US-10-146-473-72	Sequence 72, Appli
14	1878	82.1	1255	9	US-09-811-123-9	Sequence 9, Appli
15	1878	82.1	1255	9	US-09-811-115-3	Sequence 3, Appli

NAME/KEY: VARIANT
LOCATION: 356
OTHER INFORMATION: Applicants herein disclose Leu and Gln sequence variants at this
FEATURE:
NAME/KEY: VARIANT
LOCATION: 358
OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at this
FEATURE:
NAME/KEY: VARIANT
LOCATION: 361
OTHER INFORMATION: Applicants herein disclose Gln, Asp, Ala and Val sequence variants
FEATURE:
NAME/KEY: VARIANT
LOCATION: 376
OTHER INFORMATION: Applicants herein disclose Leu and Ile sequence variants at this
FEATURE:
NAME/KEY: VARIANT
LOCATION: 394
OTHER INFORMATION: Applicants herein disclose Pro and Arg sequence variants at this
FEATURE:
NAME/KEY: VARIANT
LOCATION: 404
OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
FEATURE:
NAME/KEY: VARIANT
LOCATION: 413
OTHER INFORMATION: Applicants herein disclose Asp and Asn sequence variants at this
US-10-302-663-2

Query Match 94.9%; Score 2171; DB 16; Length 419;
Best Local Similarity 95.5%; Pred. No. 6.9e-172;
Matches 400; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 1 MELAALCRWGLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQSCVVGQNL 60
DB 1 MELAALCRWGLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQSCVVGQNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVLIHNOVQVPLQRLRIVRGTLQFEDNYALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVLIHNOVQVPLQRLRIVRGTLQFEDNYALAVLNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQORNPOLCYQDTILWKDIFHKNNQLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQORNPOLCYQDTILWKDIFHKNNQLA 180
QY 181 LTLIDNTRSRACHPCSPMKSGRCWGESSEDCQSLTRTVGAGGCARCKGPLETDCCHQC 240
DB 181 LTLIDNTRSRACHPCSPMKSGRCWGESSEDCQSLTRTVGAGGCARCKGPLETDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVSGCTLVCPHLNQEVTAEDGTQRCCKSKPCARVYCYGLMEHLEHREVAVTSAN 360
DB 301 YNYLSTDVSGCTLVCPHLNQEVTAEDGTQRCCKSKPCARVYCYGLMEHLEHREVAVTSAN 360
QY 361 GPAHPVLSFLRPSWDLVSAFYSLPLAPLSPVTSVPSVSGRGPDPDAHVAVNLSRYEG 419
DB 361 XPAHPVLSFLRPSWDLVSAFYSLPLAPLSPVTSVPSVSGRGPDPDAHVAVNLSRYEG 419

RESULT 3
US-10-412-804A-6
Sequence 6, Application US/10412804A
Publication No. US20030228606A1
GENERAL INFORMATION:
APPLICANT: Jing, Shuguan
APPLICANT: Tatarewicz, Suzanna
TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
TITLE OF INVENTION: Thereof
FILE REFERENCE: 01-1624-A

CURRENT APPLICATION NUMBER: US/10/412.804A
CURRENT FILING DATE: 2003-04-11
PRIOR APPLICATION NUMBER: 60/371,912
PRIOR FILING DATE: 2002-04-11
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 6
LENGTH: 720
TYPE: PRT
ORGANISM: Homo sapiens
US-10-412-804A-6
Query Match 82.2%; Score 1879.5; DB 15; Length 720;
Best Local Similarity 76.9%; Pred. No. 2.3e-147;
Matches 367; Conservative 7; Mismatches 40; Indels 63; Gaps 7;
QY 1 MELAALCRWGLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQSCVVGQNL 60
DB 1 MELAALCRWGLLALLPPGAASCTGCTDMKRLPASPETHLDMRLHLYQSCVVGQNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVLIHNOVQVPLQRLRIVRGTLQFEDNYALAVLNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVLIHNOVQVPLQRLRIVRGTLQFEDNYALAVLNG 120
QY 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQORNPOLCYQDTILWKDIFHKNNQLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQORNPOLCYQDTILWKDIFHKNNQLA 180
QY 181 LTLIDNTRSRACHPCSPMKSGRCWGESSEDCQSLTRTVGAGGCARCKGPLETDCCHQC 240
DB 181 LTLIDNTRSRACHPCSPMKSGRCWGESSEDCQSLTRTVGAGGCARCKGPLETDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVSGCTLVCPHLNQEVTAEDGTQRCCKSKPCARVYCYGLMEHLEHREVAVTSAN 360
DB 301 YNYLSTDVSGCTLVCPHLNQEVTAEDGTQRCCKSKPCARVYCYGLMEHLEHREVAVTSAN 360
QY 361 GPAHPVLSFLRPSWDLVSAFYSLPLAPLSPVTSVPSVSGRGPDPDAHVAVNLSRYEG 419
DB 361 XPAHPVLSFLRPSWDLVSAFYSLPLAPLSPVTSVPSVSGRGPDPDAHVAVNLSRYEG 419

Query Match 82.1%; Score 1878; DB 9; Length 645;
Best Local Similarity 83.0%; Pred. No. 2.7e-147;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLLALLPFGAASSTOVCTGTDMLKRLPASPETHLDMLRLHYQCGVQVGNL 60
DB 1 MELAALCRWGLLLALLPFGAASSTOVCTGTDMLKRLPASPETHLDMLRLHYQCGVQVGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVYLIHQNVRQVPLQRLRIVRGTFQEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVYLIHQNVRQVPLQRLRIVRGTFQEDNYALAVLDNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLEILKGGVLIQPNPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLEILKGGVLIQPNPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPMKGSRGWGSESDCQSLTRTVAGGCAKCKGPLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMKGSRGWGSESDCQSLTRTVAGGCAKCKGPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR-----GTHSLPRPAVPVP 355
DB 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR-----GTHSLPRPAVPVP 355
QY 356 LRMQPG--PAHPVLSFLRPSMDLVSFAFYSLEPLAPLSPTSVP1-----SPVSVGRGPD 405
DB 356 LRMQPG--PAHPVLSFLRPSMDLVSFAFYSLEPLAPLSPTSVP1-----SPVSVGRGPD 405
QY 361 IQEFAGCKKIFGSLAFPLPESFDGDPASNT---APLQPEQLQVFTLEETITGYLYISAWPD 417
DB 361 IQEFAGCKKIFGSLAFPLPESFDGDPASNT---APLQPEQLQVFTLEETITGYLYISAWPD 417
QY 406 --PDHVAVNLSRYEG 419
DB 418 SLPDLSPFQNLQVIRG 433

RESULT 5
US-10-268-501-13
; Sequence 13, Application US/10268501
; Publication No. US20030086924A1
; GENERAL INFORMATION:
; APPLICANT: Sliwkowski, Mark X.
; TITLE OF INVENTION: Treatment with Anti-ErbB2 Antibodies
; FILE REFERENCE: P1467R2P1
; CURRENT APPLICATION NUMBER: US/10/268,501
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 09/602,812
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141,316
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 13
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-268-501-13

Query Match 82.1%; Score 1878; DB 14; Length 645;
Best Local Similarity 83.0%; Pred. No. 2.7e-147;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
QY 1 MELAALCRWGLLLALLPFGAASSTOVCTGTDMLKRLPASPETHLDMLRLHYQCGVQVGNL 60
DB 1 MELAALCRWGLLLALLPFGAASSTOVCTGTDMLKRLPASPETHLDMLRLHYQCGVQVGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVYLIHQNVRQVPLQRLRIVRGTFQEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVYLIHQNVRQVPLQRLRIVRGTFQEDNYALAVLDNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLEILKGGVLIQPNPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLEILKGGVLIQPNPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPMKGSRGWGSESDCQSLTRTVAGGCAKCKGPLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMKGSRGWGSESDCQSLTRTVAGGCAKCKGPLPTDCCHQC 240

QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR-----GTHSLPRPAVPVP 355
DB 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR-----GTHSLPRPAVPVP 355
QY 356 LRMQPG--PAHPVLSFLRPSMDLVSFAFYSLEPLAPLSPTSVP1-----SPVSVGRGPD 405
DB 356 LRMQPG--PAHPVLSFLRPSMDLVSFAFYSLEPLAPLSPTSVP1-----SPVSVGRGPD 405
QY 361 IQEFAGCKKIFGSLAFPLPESFDGDPASNT---APLQPEQLQVFTLEETITGYLYISAWPD 417
DB 361 IQEFAGCKKIFGSLAFPLPESFDGDPASNT---APLQPEQLQVFTLEETITGYLYISAWPD 417
QY 406 --PDHVAVNLSRYEG 419
DB 418 SLPDLSPFQNLQVIRG 433

RESULT 6
US-10-608-626-13
; Sequence 13, Application US/10608626
; Publication No. US20040013667A1
; GENERAL INFORMATION:
; APPLICANT: Kelsey, Stephen M.
; APPLICANT: Sliwkowski, Mark X.
; TITLE OF INVENTION: Treatment with Anti-ErbB2 Antibodies
; FILE REFERENCE: P1467R2P2
; CURRENT APPLICATION NUMBER: US/10/608,626
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: US 10/268,501
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 09/602,812
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141,316
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 13
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-608-626-13

Query Match 82.1%; Score 1878; DB 15; Length 645;
Best Local Similarity 83.0%; Pred. No. 2.7e-147;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
QY 1 MELAALCRWGLLLALLPFGAASSTOVCTGTDMLKRLPASPETHLDMLRLHYQCGVQVGNL 60
DB 1 MELAALCRWGLLLALLPFGAASSTOVCTGTDMLKRLPASPETHLDMLRLHYQCGVQVGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVYLIHQNVRQVPLQRLRIVRGTFQEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVYLIHQNVRQVPLQRLRIVRGTFQEDNYALAVLDNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLEILKGGVLIQPNPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLEILKGGVLIQPNPOLCYQDTILWKDIFHKNNOLA 180
QY 181 LTLIDTNRSRACHPCSPMKGSRGWGSESDCQSLTRTVAGGCAKCKGPLPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMKGSRGWGSESDCQSLTRTVAGGCAKCKGPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR-----GTHSLPRPAVPVP 355
DB 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCAR-----GTHSLPRPAVPVP 355
QY 356 LRMQPG--PAHPVLSFLRPSMDLVSFAFYSLEPLAPLSPTSVP1-----SPVSVGRGPD 405
DB 356 LRMQPG--PAHPVLSFLRPSMDLVSFAFYSLEPLAPLSPTSVP1-----SPVSVGRGPD 405
QY 361 IQEFAGCKKIFGSLAFPLPESFDGDPASNT---APLQPEQLQVFTLEETITGYLYISAWPD 417
DB 361 IQEFAGCKKIFGSLAFPLPESFDGDPASNT---APLQPEQLQVFTLEETITGYLYISAWPD 417

```
QY 406 --PDAHVAVNLSRYEG 419
| | | | |
Db 418 SLPDLVFNQLQVIRG 433

RESULT 7
US-09-854-356-3
; Sequence 3, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
US-09-854-356-3

Query Match 82.1%; Score 1878; DB 9; Length 653;
Best Local Similarity 83.0%; Pred. No. 2.7e-147;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMLRHLYQCCVQGNL 60
| | | | |
Db 1 MELAALCRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMLRHLYQCCVQGNL 60

QY 61 ELTYLPTNASLSFLQDIQEVQGVYLIHNVQVPLQRLRIVRGTLFEDNYALVDNG 120
| | | | |
Db 61 ELTYLPTNASLSFLQDIQEVQGVYLIHNVQVPLQRLRIVRGTLFEDNYALVDNG 120

QY 121 DPLNNTPTVTGASPGGLRELQRLSITELKGGVLIQVLPOLCYQDTILWKDIFHKNQOLA 180
| | | | |
Db 121 DPLNNTPTVTGASPGGLRELQRLSITELKGGVLIQVLPOLCYQDTILWKDIFHKNQOLA 180

QY 181 LTLIDNTRSRACHPCSPCKGSRGWGESSEDCQSLTRTVCAAGGCARCKGPLPTDCCHBOC 240
| | | | |
Db 181 LTLIDNTRSRACHPCSPCKGSRGWGESSEDCQSLTRTVCAAGGCARCKGPLPTDCCHBOC 240

QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
| | | | |
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTLVCPLNHNVTAEDGTQRCCKSKPCAR-----GTHSLPRPAAPVPP 355
| | | | |
Db 301 YNYLSTDVGSCTLVCPLNHNVTAEDGTQRCCKSKPCAR-----GTHSLPRPAAPVPP 355

QY 356 LRMQPG--PAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSVPI-----SPVSVGRGPD 405
| | | | |
Db 356 LRMQPG--PAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSVPI-----SPVSVGRGPD 405

QY 406 --PDAHVAVNLSRYEG 419
| | | | |
Db 418 SLPDLVFNQLQVIRG 433

RESULT 8
US-10-412-804A-4
; Sequence 4, Application US/10412804A
; Sequence 11, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Jings, Shuguan
; APPLICANT: Tatarewicz, Suzanna
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
; FILE REFERENCE: 01-1624-A
; CURRENT APPLICATION NUMBER: US/10/412,804A
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,912
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
```



```
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-412-804A-11

Query Match      82.1%; Score 1878; DB 15; Length 690;
Best Local Similarity 83.0%; Pred. No. 2.9e-147;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLLALLPPGAASQTGCTDMKRLPASPETHLDMLRHLHYQGCVVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASQTGCTDMKRLPASPETHLDMLRHLHYQGCVVQGNL 60

QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNNQVQVPLQRLRIVRGTLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNNQVQVPLQRLRIVRGTLFEDNYALAVLDNG 120

QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180

QY 181 LTLIDTNRSRACHPCSPMCKGSRGWGSSSEDQSLTRIVCAGGCARCKGPLEPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMCKGSRGWGSSSEDQSLTRIVCAGGCARCKGPLEPTDCCHQC 240

QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTLVCPHNOEVTAEADGTORCEKSKPCAR-----GTHSLLPRAAVPVP 355
DB 301 YNYLSTDVGSCTLVCPHNOEVTAEADGTORCEKSKPCAR-----GTHSLLPRAAVPVP 355

QY 356 LRMQPG--PAHPVLSFLRPSWDLVSATYSLPLAPLSPTSVP-----SPVSVGRGPD 405
DB 356 LRMQPG--PAHPVLSFLRPSWDLVSATYSLPLAPLSPTSVP-----SPVSVGRGPD 405

QY 361 IQEFAGCKKIFGSLAPLPESFDGDPASNT---APLQPEQLQVFTLBEITGYLYISAWPD 417
DB 361 IQEFAGCKKIFGSLAPLPESFDGDPASNT---APLQPEQLQVFTLBEITGYLYISAWPD 417

QY 406 --PDHVAVNLRYEG 419
DB 418 SLPDLVSFQNLQVIRG 433

RESULT 11
US-10-412-804A-10
; Sequence 10, Application US/10412804A
; Publication No. US20030228606A1
; GENERAL INFORMATION:
; APPLICANT: Jing, Shuguan
; APPLICANT: Tatawicz, Suzanna
; TITLE OF INVENTION: HER-2 Receptor Tyrosine Kinase Molecules and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 01-1624-A
; CURRENT APPLICATION NUMBER: US/10/412,804A
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: 60/371,912
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-412-804A-10

Query Match      82.1%; Score 1878; DB 15; Length 715;
Best Local Similarity 83.0%; Pred. No. 3.1e-147;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLLALLPPGAASQTGCTDMKRLPASPETHLDMLRHLHYQGCVVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASQTGCTDMKRLPASPETHLDMLRHLHYQGCVVQGNL 60

QY 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNNQVQVPLQRLRIVRGTLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGYVLIHNNQVQVPLQRLRIVRGTLFEDNYALAVLDNG 120

QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180
DB 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNOLA 180

QY 181 LTLIDTNRSRACHPCSPMCKGSRGWGSSSEDQSLTRIVCAGGCARCKGPLEPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPMCKGSRGWGSSSEDQSLTRIVCAGGCARCKGPLEPTDCCHQC 240

QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTLVCPHNOEVTAEADGTORCEKSKPCAR-----GTHSLLPRAAVPVP 355
DB 301 YNYLSTDVGSCTLVCPHNOEVTAEADGTORCEKSKPCAR-----GTHSLLPRAAVPVP 355

QY 356 LRMQPG--PAHPVLSFLRPSWDLVSATYSLPLAPLSPTSVP-----SPVSVGRGPD 405
DB 356 LRMQPG--PAHPVLSFLRPSWDLVSATYSLPLAPLSPTSVP-----SPVSVGRGPD 405

QY 361 IQEFAGCKKIFGSLAPLPESFDGDPASNT---APLQPEQLQVFTLBEITGYLYISAWPD 417
DB 361 IQEFAGCKKIFGSLAPLPESFDGDPASNT---APLQPEQLQVFTLBEITGYLYISAWPD 417

QY 406 --PDHVAVNLRYEG 419
DB 418 SLPDLVSFQNLQVIRG 433

RESULT 10
US-09-854-356-7
; Sequence 7, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 7
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein
; OTHER INFORMATION: of ECD and delta PD of human HER-2/neu
US-09-854-356-7

Query Match      82.1%; Score 1878; DB 9; Length 712;
Best Local Similarity 83.0%; Pred. No. 3e-147;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
```

Db 181 LTLIDNRSRACHPCSPMCKGRCWGESSEDCQSLTRTVAGGCARCKGPLEPTDCCHQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 355
Db 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 360
Qy 356 LRMQPG--PAHPVLSFLRPSMDLVSFYSPLAPLSPTSVP1-----SPVSVGRGPD 405
Db 361 IQEAGCKKIFGSLAFPESEFDGDPASNT---APLQPEQLQVFTLEETITGYLISAWPD 417
Qy 406 --PDHVAVNLSRYEG 419
Db 418 SLPLSVFQNLQVIRG 433
RESULT 12
US-09-854-356-6
; Sequence 6, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009A10PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: fusion protein
; OTHER INFORMATION: of ECD and PD of human HER-2/neu
US-09-854-356-6
Query Match 82.1%; Score 1878; DB 9; Length 919;
Best Local Similarity 83.0%; Pred. No. 4.2e-147;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
Qy 1 MELAALCRWGLLALLPPGAASQVCTGTDMLRLPASPEHLDMLRHLRYGCVVQGNL 60
Db 1 MELAALCRWGLLALLPPGAASQVCTGTDMLRLPASPEHLDMLRHLRYGCVVQGNL 60
Qy 61 ELTYLPTNASLSFLQDIQEVQGVLIHNRQVPLQRLRIVRGTLQFEDNVALAVLNG 120
Db 61 ELTYLPTNASLSFLQDIQEVQGVLIHNRQVPLQRLRIVRGTLQFEDNVALAVLNG 120
Qy 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQORNPOLCVQDTILWKDIFHKNQOLA 180
Db 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQORNPOLCVQDTILWKDIFHKNQOLA 180
Qy 181 LTLIDNRSRACHPCSPMCKGRCWGESSEDCQSLTRTVAGGCARCKGPLEPTDCCHQC 240
Db 181 LTLIDNRSRACHPCSPMCKGRCWGESSEDCQSLTRTVAGGCARCKGPLEPTDCCHQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 355
Db 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 360
Qy 356 LRMQPG--PAHPVLSFLRPSMDLVSFYSPLAPLSPTSVP1-----SPVSVGRGPD 405
Db 361 IQEAGCKKIFGSLAFPESEFDGDPASNT---APLQPEQLQVFTLEETITGYLISAWPD 417
Qy 406 --PDHVAVNLSRYEG 419
Db 418 SLPLSVFQNLQVIRG 433

Db 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 360
Qy 356 LRMQPG--PAHPVLSFLRPSMDLVSFYSPLAPLSPTSVP1-----SPVSVGRGPD 405
Db 361 IQEAGCKKIFGSLAFPESEFDGDPASNT---APLQPEQLQVFTLEETITGYLISAWPD 417
Qy 406 --PDHVAVNLSRYEG 419
Db 418 SLPLSVFQNLQVIRG 433
RESULT 13
US-10-146-473-72
; Sequence 72, Application US/10146473
; Publication No. US2003010888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: 100461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72
; LENGTH: 1253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-72
Query Match 82.1%; Score 1878; DB 14; Length 1253;
Best Local Similarity 83.0%; Pred. No. 6.2e-147;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
Qy 1 MELAALCRWGLLALLPPGAASQVCTGTDMLRLPASPEHLDMLRHLRYGCVVQGNL 60
Db 1 MELAALCRWGLLALLPPGAASQVCTGTDMLRLPASPEHLDMLRHLRYGCVVQGNL 60
Qy 61 ELTYLPTNASLSFLQDIQEVQGVLIHNRQVPLQRLRIVRGTLQFEDNVALAVLNG 120
Db 61 ELTYLPTNASLSFLQDIQEVQGVLIHNRQVPLQRLRIVRGTLQFEDNVALAVLNG 120
Qy 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQORNPOLCVQDTILWKDIFHKNQOLA 180
Db 121 DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQORNPOLCVQDTILWKDIFHKNQOLA 180
Qy 181 LTLIDNRSRACHPCSPMCKGRCWGESSEDCQSLTRTVAGGCARCKGPLEPTDCCHQC 240
Db 181 LTLIDNRSRACHPCSPMCKGRCWGESSEDCQSLTRTVAGGCARCKGPLEPTDCCHQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 355
Db 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 360
Qy 356 LRMQPG--PAHPVLSFLRPSMDLVSFYSPLAPLSPTSVP1-----SPVSVGRGPD 405
Db 361 IQEAGCKKIFGSLAFPESEFDGDPASNT---APLQPEQLQVFTLEETITGYLISAWPD 417
Qy 406 --PDHVAVNLSRYEG 419
Db 418 SLPLSVFQNLQVIRG 433

RESULT 14

US-09-811-123-9
; Sequence 9, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
; FILE REFERENCE: GENENT.073A2
; CURRENT APPLICATION NUMBER: US/09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-123-9

Query Match 82.1%; Score 1878; DB 9; Length 1255;

Best Local Similarity 83.0%; Pred. No. 6.2e-147; Indels 20; Gaps 5;
Matches 362; Conservative 9; Mismatches 45;

QY	1	MELAAACRWGLLLALLPPGAAS	TQVCTG	TMKRLR	PASPETH	LDMLRHL	YQCCVQGNL	60
DB	1	MELAAACRWGLLLALLPPGAAS	TQVCTG	TMKRLR	PASPETH	LDMLRHL	YQCCVQGNL	60
QY	61	ELTYLPTNASL	SFLQDIQ	VGQVYLI	AHNOVQ	PLQRLR	IVRG	QVLPEDNYALAVLNG 120
DB	61	ELTYLPTNASL	SFLQDIQ	VGQVYLI	AHNOVQ	PLQRLR	IVRG	QVLPEDNYALAVLNG 120
QY	121	DPLNNTT	PVTGAS	PGGLREL	QLRSL	TEILK	GGVLIQ	RNPOLCYQDTILWKDIFHKNQLA 180
DB	121	DPLNNTT	PVTGAS	PGGLREL	QLRSL	TEILK	GGVLIQ	RNPOLCYQDTILWKDIFHKNQLA 180
QY	181	LTLLD	TNRSRACH	PCSPCK	GRGRCW	GSESSDC	QSLTR	TVCAAGCARCKGPLETDCCHQC 240
DB	181	LTLLD	TNRSRACH	PCSPCK	GRGRCW	GSESSDC	QSLTR	TVCAAGCARCKGPLETDCCHQC 240
QY	241	AGCTG	PKHSDCL	ACLHFN	HSGI	CELHCP	ALVTYNTD	TFESMPNPEGRYTFGASCVTACP 300
DB	241	AGCTG	PKHSDCL	ACLHFN	HSGI	CELHCP	ALVTYNTD	TFESMPNPEGRYTFGASCVTACP 300
QY	301	YNYL	STDVGS	CTLVCP	LHNOE	VTAE	DGTQ	RCCKSPCARVCYGLGMEHLREVRAVTSAN 360
DB	301	YNYL	STDVGS	CTLVCP	LHNOE	VTAE	DGTQ	RCCKSPCARVCYGLGMEHLREVRAVTSAN 360
QY	356	LRMQPG	--PAHP	VLSFLR	PSWDL	VS	AFYSL	PLAPLSPTSVPI-----SPVSVGRGPD 405
DB	356	LRMQPG	--PAHP	VLSFLR	PSWDL	VS	AFYSL	PLAPLSPTSVPI-----SPVSVGRGPD 405
QY	406	--PD	AVAVNL	SR	YEG	419		
DB	406	--PD	AVAVNL	SR	YEG	419		
QY	418	SLP	DL	SVFQ	NLQ	VIRG	433	
DB	418	SLP	DL	SVFQ	NLQ	VIRG	433	

RESULT 15

US-09-811-115-3
; Sequence 3, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A

; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3

Query Match 82.1%; Score 1878; DB 9; Length 1255;

Best Local Similarity 83.0%; Pred. No. 6.2e-147; Indels 20; Gaps 5;
Matches 362; Conservative 9; Mismatches 45;

QY	1	MELAAACRWGLLLALLPPGAAS	TQVCTG	TMKRLR	PASPETH	LDMLRHL	YQCCVQGNL	60
DB	1	MELAAACRWGLLLALLPPGAAS	TQVCTG	TMKRLR	PASPETH	LDMLRHL	YQCCVQGNL	60
QY	61	ELTYLPTNASL	SFLQDIQ	VGQVYLI	AHNOVQ	PLQRLR	IVRG	QVLPEDNYALAVLNG 120
DB	61	ELTYLPTNASL	SFLQDIQ	VGQVYLI	AHNOVQ	PLQRLR	IVRG	QVLPEDNYALAVLNG 120
QY	121	DPLNNTT	PVTGAS	PGGLREL	QLRSL	TEILK	GGVLIQ	RNPOLCYQDTILWKDIFHKNQLA 180
DB	121	DPLNNTT	PVTGAS	PGGLREL	QLRSL	TEILK	GGVLIQ	RNPOLCYQDTILWKDIFHKNQLA 180
QY	181	LTLLD	TNRSRACH	PCSPCK	GRGRCW	GSESSDC	QSLTR	TVCAAGCARCKGPLETDCCHQC 240
DB	181	LTLLD	TNRSRACH	PCSPCK	GRGRCW	GSESSDC	QSLTR	TVCAAGCARCKGPLETDCCHQC 240
QY	241	AGCTG	PKHSDCL	ACLHFN	HSGI	CELHCP	ALVTYNTD	TFESMPNPEGRYTFGASCVTACP 300
DB	241	AGCTG	PKHSDCL	ACLHFN	HSGI	CELHCP	ALVTYNTD	TFESMPNPEGRYTFGASCVTACP 300
QY	301	YNYL	STDVGS	CTLVCP	LHNOE	VTAE	DGTQ	RCCKSPCARVCYGLGMEHLREVRAVTSAN 360
DB	301	YNYL	STDVGS	CTLVCP	LHNOE	VTAE	DGTQ	RCCKSPCARVCYGLGMEHLREVRAVTSAN 360
QY	356	LRMQPG	--PAHP	VLSFLR	PSWDL	VS	AFYSL	PLAPLSPTSVPI-----SPVSVGRGPD 405
DB	356	LRMQPG	--PAHP	VLSFLR	PSWDL	VS	AFYSL	PLAPLSPTSVPI-----SPVSVGRGPD 405
QY	406	--PD	AVAVNL	SR	YEG	419		
DB	406	--PD	AVAVNL	SR	YEG	419		
QY	418	SLP	DL	SVFQ	NLQ	VIRG	433	
DB	418	SLP	DL	SVFQ	NLQ	VIRG	433	

RESULT 16

US-09-769-508-2
; Sequence 2, Application US/09769508
; Patent No. US20020155527A1
; GENERAL INFORMATION:
; APPLICANT: STUART, SUSAN G.
; APPLICANT: MONAHAN, JOHN J.
; APPLICANT: LANGTON, BEATRICE CLAUDIA
; APPLICANT: HANCOCK, MIRIAM E.C.
; APPLICANT: CHAO, LORRINE A.
; APPLICANT: BLUFORD, PETER
; TITLE OF INVENTION: C-ERBB-2
; FILE REFERENCE: BEBIO-111-C1
; CURRENT APPLICATION NUMBER: US/09/769,508
; CURRENT FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-508-2

Query Match 82.1%; Score 1878; DB 9; Length 1255;

Best Local Similarity 83.0%; Pred. No. 6.2e-147; Indels 20; Gaps 5;
Matches 362; Conservative 9; Mismatches 45;

QY 1 MELAALCRWGLLLALLPPGAASQTCTGDKMLRLPASPETHLMLRLHLYGCGVQGNL 60
DQ 1 MELAALCRWGLLLALLPPGAASQTCTGDKMLRLPASPETHLMLRLHLYGCGVQGNL 60

QY 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNOVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
DQ 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNOVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120

QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
DQ 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180

QY 181 LTLIDTNRSRACHPCSPCKGSRGWESSEDCQSLTRVTCAGGCARCKGPLPTDCCHQC 240
DQ 181 LTLIDTNRSRACHPCSPCKGSRGWESSEDCQSLTRVTCAGGCARCKGPLPTDCCHQC 240

QY 241 AAGCTGPKHSDCLACILHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DQ 241 AAGCTGPKHSDCLACILHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTLVCPFLHNOEVTAEADGTQRCCKSKPCAR-----GTHSLPRPAAVPP 355
DQ 301 YNYLSTDVGSCTLVCPFLHNOEVTAEADGTQRCCKSKPCAR-----GTHSLPRPAAVPP 355

QY 356 LEMQPG--PAHPVLSFLRPSMDLVSAFVSLPLAPLSPTSVPI-----SPVSVGRGPD 405
DQ 356 LEMQPG--PAHPVLSFLRPSMDLVSAFVSLPLAPLSPTSVPI-----SPVSVGRGPD 405

QY 406 ---PDHVAVNLRYEG 419
DQ 406 ---PDHVAVNLRYEG 419

QY 418 SLPDLVSFQNLQVIRG 433
DQ 418 SLPDLVSFQNLQVIRG 433

RESULT 17
US-09-854-356-1
; Sequence 1, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1993-01-29
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)...(653)
; OTHER INFORMATION: extracellular domain (ECD)
; NAME/KEY: DOMAIN
; LOCATION: (676)...(1255)
; OTHER INFORMATION: intracellular domain (ICD)
; NAME/KEY: DOMAIN
; LOCATION: (990)...(1255)
; OTHER INFORMATION: phosphorylation domain (PD)
; NAME/KEY: DOMAIN
; LOCATION: (990)...(1048)
; OTHER INFORMATION: fragment of the phosphorylation domain, preferred

OTHER INFORMATION: portion (delta PD)
US-09-854-356-1

Query Match 82.1%; Score 1878; DB 9; Length 1255;
Best Local Similarity 83.0%; Pred. No. 6.2e-147; Indels 20; Gaps 5;
Matches 362; Conservative 9; Mismatches 45;

QY 1 MELAALCRWGLLLALLPPGAASQTCTGDKMLRLPASPETHLMLRLHLYGCGVQGNL 60
DQ 1 MELAALCRWGLLLALLPPGAASQTCTGDKMLRLPASPETHLMLRLHLYGCGVQGNL 60

QY 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNOVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
DQ 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNOVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120

QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
DQ 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180

QY 181 LTLIDTNRSRACHPCSPCKGSRGWESSEDCQSLTRVTCAGGCARCKGPLPTDCCHQC 240
DQ 181 LTLIDTNRSRACHPCSPCKGSRGWESSEDCQSLTRVTCAGGCARCKGPLPTDCCHQC 240

QY 241 AAGCTGPKHSDCLACILHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DQ 241 AAGCTGPKHSDCLACILHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCTLVCPFLHNOEVTAEADGTQRCCKSKPCAR-----GTHSLPRPAAVPP 355
DQ 301 YNYLSTDVGSCTLVCPFLHNOEVTAEADGTQRCCKSKPCAR-----GTHSLPRPAAVPP 355

QY 356 LEMQPG--PAHPVLSFLRPSMDLVSAFVSLPLAPLSPTSVPI-----SPVSVGRGPD 405
DQ 356 LEMQPG--PAHPVLSFLRPSMDLVSAFVSLPLAPLSPTSVPI-----SPVSVGRGPD 405

QY 406 ---PDHVAVNLRYEG 419
DQ 406 ---PDHVAVNLRYEG 419

QY 418 SLPDLVSFQNLQVIRG 433
DQ 418 SLPDLVSFQNLQVIRG 433

RESULT 18
US-09-930-125-2
; Sequence 2, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedwick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-930-125-2

Query Match 82.1%; Score 1878; DB 9; Length 1255;
Best Local Similarity 83.0%; Pred. No. 6.2e-147; Indels 20; Gaps 5;
Matches 362; Conservative 9; Mismatches 45;

QY 1 MELAALCRWGLLLALLPPGAASQTCTGDKMLRLPASPETHLMLRLHLYGCGVQGNL 60
DQ 1 MELAALCRWGLLLALLPPGAASQTCTGDKMLRLPASPETHLMLRLHLYGCGVQGNL 60

```
QY 61 ELTYLPTNASLSFLQDIOEVQGYVLI AHNOVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIOEVQGYVLI AHNOVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
QY 121 DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNOLA 180
Db 121 DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNOLA 180
QY 181 LTLIDTNRSRACHPCSPMKGSRGCGWGESSEDCQSLTRTVACGGCARCKGPLPTDCCHEQC 240
Db 181 LTLIDTNRSRACHPCSPMKGSRGCGWGESSEDCQSLTRTVACGGCARCKGPLPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNOEVTAEQTCRCKSKPCAR-----GTHSLPRPAAVVP 355
Db 301 YNYLSTDVGSCTLVCPHNOEVTAEQTCRCKSKPCAR-----GTHSLPRPAAVVP 355
QY 356 LRMQPG--PAHPVLSFLRPSWDLVSFSLPLAFLSPTSVP-----SPVSVGRGPD 405
Db 356 LRMQPG--PAHPVLSFLRPSWDLVSFSLPLAFLSPTSVP-----SPVSVGRGPD 405
QY 361 IQEPAGCKKIFGSLAFPLPESFDGPASNT---APLQPEQLQVFTLEETGYLYISAWPD 417
Db 361 IQEPAGCKKIFGSLAFPLPESFDGPASNT---APLQPEQLQVFTLEETGYLYISAWPD 417
QY 406 --PDHVAVNLSRYEG 419
Db 418 SLPDLVSFQNLQVIRG 433
RESULT 19
US-09-441-411-6
; Sequence 6, Application US/09441411
; Publication No. US20030008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033-409
; CURRENT APPLICATION NUMBER: US/09/441,411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-411-6
Query Match 82.1%; Score 1878; DB 10; Length 1255;
Best Local Similarity 83.0%; Pred. No. 6.2e-147;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
QY 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPETHDMLRLHLYQGCVVQGNL 60
Db 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPETHDMLRLHLYQGCVVQGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGYVLI AHNOVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIOEVQGYVLI AHNOVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
QY 121 DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNOLA 180
Db 121 DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNOLA 180
QY 181 LTLIDTNRSRACHPCSPMKGSRGCGWGESSEDCQSLTRTVACGGCARCKGPLPTDCCHEQC 240
Db 181 LTLIDTNRSRACHPCSPMKGSRGCGWGESSEDCQSLTRTVACGGCARCKGPLPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNOEVTAEQTCRCKSKPCAR-----GTHSLPRPAAVVP 355
Db 301 YNYLSTDVGSCTLVCPHNOEVTAEQTCRCKSKPCAR-----GTHSLPRPAAVVP 355
QY 356 LRMQPG--PAHPVLSFLRPSWDLVSFSLPLAFLSPTSVP-----SPVSVGRGPD 405
Db 356 LRMQPG--PAHPVLSFLRPSWDLVSFSLPLAFLSPTSVP-----SPVSVGRGPD 405
QY 361 IQEPAGCKKIFGSLAFPLPESFDGPASNT---APLQPEQLQVFTLEETGYLYISAWPD 417
Db 361 IQEPAGCKKIFGSLAFPLPESFDGPASNT---APLQPEQLQVFTLEETGYLYISAWPD 417
QY 406 --PDHVAVNLSRYEG 419
Db 418 SLPDLVSFQNLQVIRG 433
RESULT 21
US-10-469-162-3
; Sequence 3, Application US/10469162
; Publication No. US20040052811A1
```

```
QY 301 YNYLSTDVGSCTLVCPHNOEVTAEQTCRCKSKPCAR-----GTHSLPRPAAVVP 355
Db 301 YNYLSTDVGSCTLVCPHNOEVTAEQTCRCKSKPCAR-----GTHSLPRPAAVVP 355
QY 356 LRMQPG--PAHPVLSFLRPSWDLVSFSLPLAFLSPTSVP-----SPVSVGRGPD 405
Db 356 LRMQPG--PAHPVLSFLRPSWDLVSFSLPLAFLSPTSVP-----SPVSVGRGPD 405
QY 361 IQEPAGCKKIFGSLAFPLPESFDGPASNT---APLQPEQLQVFTLEETGYLYISAWPD 417
Db 361 IQEPAGCKKIFGSLAFPLPESFDGPASNT---APLQPEQLQVFTLEETGYLYISAWPD 417
QY 406 --PDHVAVNLSRYEG 419
Db 418 SLPDLVSFQNLQVIRG 433
RESULT 20
US-09-984-092-4
; Sequence 4, Application US/09984092
; Publication No. US20040037840A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL THERAPEUTIC VACCINE FORMULATIONS
; FILE REFERENCE: F1011PC00
; CURRENT APPLICATION NUMBER: US/09/984,092
; CURRENT FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-092-4
Query Match 82.1%; Score 1878; DB 12; Length 1255;
Best Local Similarity 83.0%; Pred. No. 6.2e-147;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
QY 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPETHDMLRLHLYQGCVVQGNL 60
Db 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPETHDMLRLHLYQGCVVQGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGYVLI AHNOVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIOEVQGYVLI AHNOVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
QY 121 DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNOLA 180
Db 121 DPLNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNOLA 180
QY 181 LTLIDTNRSRACHPCSPMKGSRGCGWGESSEDCQSLTRTVACGGCARCKGPLPTDCCHEQC 240
Db 181 LTLIDTNRSRACHPCSPMKGSRGCGWGESSEDCQSLTRTVACGGCARCKGPLPTDCCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNOEVTAEQTCRCKSKPCAR-----GTHSLPRPAAVVP 355
Db 301 YNYLSTDVGSCTLVCPHNOEVTAEQTCRCKSKPCAR-----GTHSLPRPAAVVP 355
QY 356 LRMQPG--PAHPVLSFLRPSWDLVSFSLPLAFLSPTSVP-----SPVSVGRGPD 405
Db 356 LRMQPG--PAHPVLSFLRPSWDLVSFSLPLAFLSPTSVP-----SPVSVGRGPD 405
QY 361 IQEPAGCKKIFGSLAFPLPESFDGPASNT---APLQPEQLQVFTLEETGYLYISAWPD 417
Db 361 IQEPAGCKKIFGSLAFPLPESFDGPASNT---APLQPEQLQVFTLEETGYLYISAWPD 417
QY 406 --PDHVAVNLSRYEG 419
Db 418 SLPDLVSFQNLQVIRG 433
RESULT 21
US-10-469-162-3
; Sequence 3, Application US/10469162
; Publication No. US20040052811A1
```

GENERAL INFORMATION:
; APPLICANT: Zielinski, Christoph
; APPLICANT: Behnkeberger, Hubert
; APPLICANT: Breiteneder, Helmo
; APPLICANT: Jensen-Jarolim, Erika
; APPLICANT: Scheiner, Otto
; TITLE OF INVENTION: Vaccines Against Cancerous Diseases Associated With the HER-2/neu
; TITLE OF INVENTION: oncogene
; FILE REFERENCE: K 38 132/3yv
; CURRENT APPLICATION NUMBER: US/10/469,162
; PRIOR FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: PCT/EP02/02111
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: EP 01104943.4
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(675)
; OTHER INFORMATION: Extracellular Domain
US-10-469-162-3

Query Match 82.1%; Score 1878; DB 12; Length 1255;
Best Local Similarity 83.0%; Pred. No. 6.2e-147; Indels 20; Gaps 5;
Matches 362; Conservative 9; Mismatches 45;
QY 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLKRLPASPETHLDMLRLHYQGCQVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLKRLPASPETHLDMLRLHYQGCQVQGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGVYLIHNRQVPLQRLIRVRGTQLPEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGVYLIHNRQVPLQRLIRVRGTQLPEDNYALAVLDNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPQLCYQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPQLCYQDTILWKDIFHKNQOLA 180
QY 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVGAGGCARCKGPLTDCCHQOC 240
DB 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVGAGGCARCKGPLTDCCHQOC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNRQVETADGTQRCCKSKPCAR-----GTHSLPRPAAPVP 355
DB 301 YNYLSTDVGSCTLVCPHNRQVETADGTQRCCKSKPCAR-----GTHSLPRPAAPVP 355
QY 356 LRMQPG--PAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSVP-----SPVSVGRGPD 405
DB 361 IQEFAKKIFGSLAFPSFDGDPASNT---APLQPEQLQVFTLEBTGLYISAWPD 417
QY 406 --PDHVAVNLSRYEG 419
DB 418 SLPDLVSFQNLQVIRG 433

RESULT 22
US-10-253-286-553
; Sequence 553, Application US/10253286
; Publication No. US20040058881A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: II-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015

CURRENT APPLICATION NUMBER: US/10/253,286
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 553
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-253-286-553
Query Match 82.1%; Score 1878; DB 12; Length 1255;
Best Local Similarity 83.0%; Pred. No. 6.2e-147; Indels 20; Gaps 5;
Matches 362; Conservative 9; Mismatches 45;
QY 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLKRLPASPETHLDMLRLHYQGCQVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASQVCTGTDMLKRLPASPETHLDMLRLHYQGCQVQGNL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGVYLIHNRQVPLQRLIRVRGTQLPEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGVYLIHNRQVPLQRLIRVRGTQLPEDNYALAVLDNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPQLCYQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPQLCYQDTILWKDIFHKNQOLA 180
QY 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVGAGGCARCKGPLTDCCHQOC 240
DB 181 LTLIDNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVGAGGCARCKGPLTDCCHQOC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHNRQVETADGTQRCCKSKPCAR-----GTHSLPRPAAPVP 355
DB 301 YNYLSTDVGSCTLVCPHNRQVETADGTQRCCKSKPCAR-----GTHSLPRPAAPVP 355
QY 356 LRMQPG--PAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSVP-----SPVSVGRGPD 405
DB 361 IQEFAKKIFGSLAFPSFDGDPASNT---APLQPEQLQVFTLEBTGLYISAWPD 417
QY 406 --PDHVAVNLSRYEG 419
DB 418 SLPDLVSFQNLQVIRG 433

RESULT 23
US-09-765-973-2
; Sequence 2, Application US/09765973
; Publication No. US20020039573A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Hand-Zimmermann, Susan
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR PREVENTION AND
; TITLE OF INVENTION: TREATMENT OF HER-2/neu ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.496
; CURRENT APPLICATION NUMBER: US/09/765,973
; CURRENT FILING DATE: 2001-01-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-765-973-2

Query Match 82.1%; Score 1878; DB 12; Length 1255;
Best Local Similarity 83.0%; Pred. No. 6.2e-147;

Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
Qy 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLPASPETHLDMLRHLYQGQVVOGNL 60
Db 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLPASPETHLDMLRHLYQGQVVOGNL 60
Qy 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
Qy 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNNOLA 180
Db 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNNOLA 180
Qy 181 LTLIDNRSRACHPCSPMKSGRCWGSSEDCQSLTRTVCAAGCCKGKGLPTDCCHQC 240
Db 181 LTLIDNRSRACHPCSPMKSGRCWGSSEDCQSLTRTVCAAGCCKGKGLPTDCCHQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMNPGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMNPGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVYNTDTFESMNPGRYTFGASCVTACP 355
Db 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVYNTDTFESMNPGRYTFGASCVTACP 355
Qy 356 LRMQPG--PAHPVLSFLRPSWDLVSFYSPLAPLSPTSVP-----GTHSLPRPAVPVP 405
Db 356 LRMQPG--PAHPVLSFLRPSWDLVSFYSPLAPLSPTSVP-----GTHSLPRPAVPVP 405
Qy 361 IQBFAGCKIFGSLAFPLPESFDGDPASNT---APLQPEQLQVFETLEETGVLVISAWPD 417
Db 361 IQBFAGCKIFGSLAFPLPESFDGDPASNT---APLQPEQLQVFETLEETGVLVISAWPD 417
Qy 406 --PDAHVNLSRYEG 419
Db 406 --PDAHVNLSRYEG 419
Qy 418 SLPDLVSFQNLQVIRG 433
Db 418 SLPDLVSFQNLQVIRG 433

RESULT 24
US-10-418-027-3
; Sequence 3, Application US/10418027
; Publication No. US2003024467A1
; GENERAL INFORMATION:
; APPLICANT: Osborne, C. Kent
; APPLICANT: Schiffr, Rachel
; APPLICANT: Bardou, Valerie
; APPLICANT: Hilsenbeck, Susan
; APPLICANT: Clark, Gary
; APPLICANT: Wong, Jiemin
; APPLICANT: Chamness, Gary
; APPLICANT: Hopp, Torsten
; TITLE OF INVENTION: AIB 1 as a prognostic marker and predictor of endocrine therapy
; FILE REFERENCE: HO-P02396US1
; CURRENT APPLICATION NUMBER: US/10/418,027
; PRIOR FILING DATE: 2003-04-17
; PRIOR FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Human
US-10-418-027-3

Query Match 82.1%; Score 1878; DB 12; Length 1255;
Best Local Similarity 83.0%; Pred. No. 6.2e-147;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
Qy 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLPASPETHLDMLRHLYQGQVVOGNL 60
Db 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLPASPETHLDMLRHLYQGQVVOGNL 60
Qy 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120

Db 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
Qy 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNNOLA 180
Db 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNNOLA 180
Qy 181 LTLIDNRSRACHPCSPMKSGRCWGSSEDCQSLTRTVCAAGCCKGKGLPTDCCHQC 240
Db 181 LTLIDNRSRACHPCSPMKSGRCWGSSEDCQSLTRTVCAAGCCKGKGLPTDCCHQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMNPGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMNPGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVYNTDTFESMNPGRYTFGASCVTACP 355
Db 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVYNTDTFESMNPGRYTFGASCVTACP 355
Qy 356 LRMQPG--PAHPVLSFLRPSWDLVSFYSPLAPLSPTSVP-----GTHSLPRPAVPVP 405
Db 356 LRMQPG--PAHPVLSFLRPSWDLVSFYSPLAPLSPTSVP-----GTHSLPRPAVPVP 405
Qy 361 IQBFAGCKIFGSLAFPLPESFDGDPASNT---APLQPEQLQVFETLEETGVLVISAWPD 417
Db 361 IQBFAGCKIFGSLAFPLPESFDGDPASNT---APLQPEQLQVFETLEETGVLVISAWPD 417
Qy 406 --PDAHVNLSRYEG 419
Db 406 --PDAHVNLSRYEG 419
Qy 418 SLPDLVSFQNLQVIRG 433
Db 418 SLPDLVSFQNLQVIRG 433

RESULT 25
US-10-207-655-45
; Sequence 45, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 45
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-45

Query Match 82.1%; Score 1878; DB 14; Length 1255;
Best Local Similarity 83.0%; Pred. No. 6.2e-147;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
Qy 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLPASPETHLDMLRHLYQGQVVOGNL 60
Db 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLPASPETHLDMLRHLYQGQVVOGNL 60
Qy 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIOEVQGYVLIHNVQVPLQRLRIVRGTLQFEDNYALAVLDNG 120
Qy 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNNOLA 180
Db 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQORNPOLCYQDTILWKDIFHKNNOLA 180
Qy 181 LTLIDNRSRACHPCSPMKSGRCWGSSEDCQSLTRTVCAAGCCKGKGLPTDCCHQC 240
Db 181 LTLIDNRSRACHPCSPMKSGRCWGSSEDCQSLTRTVCAAGCCKGKGLPTDCCHQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMNPGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMNPGRYTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVYNTDTFESMNPGRYTFGASCVTACP 355
Db 301 YNYLSTDVGSCTLVCPHNSGICELHCPALVYNTDTFESMNPGRYTFGASCVTACP 355

```

356 LRMQEG--PAHPVLSFLRPSNDLIVASYSFLAPLSPTSVPI-----SPVSVGRGPD 405
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
361 IQEAGCKKIKGSLAFUFESDGDPAST---APLQPEQLQVFTLEIRITGYLY-SAMPD 417
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
406 --PDHVAVNLRSRYEG 419
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
418 SLPLDSVPQNQLQVRG 433
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

```

RESULT 26
US-10-177-293-126
: Sequence 126, Application US/10177293
: Publication No. US20030124128A1
: GENERAL INFORMATION:
: APPLICANT: Lillie, James
: APPLICANT: Glatt, Karen
: APPLICANT: Zhao, Xumei
: APPLICANT: Gannavarpu, Manjula
: APPLICANT: Kamatkar, Shubhangi
: APPLICANT: Mertens, Maureen
: APPLICANT: Myer, Vic
: APPLICANT: Wang, Youzhen
: APPLICANT: Xu, Yongyao
: APPLICANT: Hoersch, Sebastian
: APPLICANT: Monahan, John
: APPLICANT: Meyers, Rachel E.
: APPLICANT: Bast Jr., Robert C.
: APPLICANT: Hortobagyi, Gabriel N.
: APPLICANT: Pusztai, Lajos
: APPLICANT: Meric, Funda
: APPLICANT: Sahin, Aysegul
: APPLICANT: Mills, Gordon B.
: TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
: FILE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
: FILE REFERENCE: MRI-038
: CURRENT APPLICATION NUMBER: US/10/177,293
: PRIOR FILING DATE: 2002-06-21
: PRIOR APPLICATION NUMBER: US 60/299,887
: PRIOR FILING DATE: 2001-06-21
: PRIOR APPLICATION NUMBER: US 60/301,572
: PRIOR FILING DATE: 2001-06-27
: PRIOR APPLICATION NUMBER: US 60/306,501
: PRIOR FILING DATE: 2001-07-18
: PRIOR APPLICATION NUMBER: US 60/325,002
: PRIOR FILING DATE: 2001-09-25
: PRIOR APPLICATION NUMBER: US 60/362,585
: PRIOR FILING DATE: 2002-03-05
: PRIOR APPLICATION NUMBER: US 60/xxx,xxx
: PRIOR FILING DATE: 2002-05-14
: NUMBER OF SEQ ID NOS: 506
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 126
: LENGTH: 1255
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-177-293-126

```

Query Match	82.1%	Score 1878;	DB 14;	Length 1255;
Best Local Similarity	83.0%;	Pred. No. 6.2e-147;		
Matches 362;	Conservative	9; Mismatches 45;	Indels	20; Gaps 5;
1	MELAAALCRWGLLALLPPGNAASQVCTGTDMLRLPASPETHLDMLRHLYGCCQVQGNL	60		
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51	ELTYLPTNASLSFLQDIQIEVQGVVLIAHNQVQVPLQRLRIVRGTLFEDNYALAVLDNG	120		
52				
53				
54				
55				
56				
57				
58				
59				
60				
61	ELTYLPTNASLSFLQDIQIEVQGVVLIAHNQVQVPLQRLRIVRGTLFEDNYALAVLDNG	120		
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				
84				
85				
86				
87				
88				
89				
90				
91				
92				
93				
94				
95				
96				
97				
98				
99				
100				


```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-322-892-4

Query Match      82.1%; Score 1378; DB 14; Length 1255;
Best Local Similarity 83.0%; Pred. No. 6.2e-147;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;

QY 1 MELAALCRWGLLLALLPPGAASQVCTGTDMKRLPASPETHLDMLRHLHYQCQVQGNL 60
Db 1 MELAALCRWGLLLALLPPGAASQVCTGTDMKRLPASPETHLDMLRHLHYQCQVQGNL 60

QY 61 ELTYLPTNASLSFLQDIQEVQGVLIAHNOVQVPLQRLRIVRGTQLPEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIQEVQGVLIAHNOVQVPLQRLRIVRGTQLPEDNYALAVLDNG 120

QY 121 DPLNNTTPTVGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180
Db 121 DPLNNTTPTVGASPGGLRELQRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQOLA 180

QY 181 LTLIDTNRSRACHPCSPMCKGRCWGESSEDCQSLTRTYCAGSCARCKGPLPTDCCHQOC 240
Db 181 LTLIDTNRSRACHPCSPMCKGRCWGESSEDCQSLTRTYCAGSCARCKGPLPTDCCHQOC 240

QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300

QY 301 YNYLSTDVGSCITVCPHLNHOEVTAEQTCQCKSKPCAR-----GTHSLLPRAVPVP 355
Db 301 YNYLSTDVGSCITVCPHLNHOEVTAEQTCQCKSKPCAR-----GTHSLLPRAVPVP 355

QY 356 LRMQPG--PAHPVLSFLRPSWDLVAFYSLPLAPLSPTSVPI-----SPVSVGRGPD 405
Db 356 LRMQPG--PAHPVLSFLRPSWDLVAFYSLPLAPLSPTSVPI-----SPVSVGRGPD 405

QY 406 --PDAHVAVNLSRYEG 419
Db 406 --PDAHVAVNLSRYEG 419

QY 418 SLFDLSVFQNLQVIRG 433
Db 418 SLFDLSVFQNLQVIRG 433
```

Search completed: July 4, 2004, 04:29:42
Job time : 85.9779 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2004, 04:17:04 ; Search time 29.4478 Seconds
 (without alignments)
 1368.668 Million cell updates/sec

Title: US-09-506-079H-12
 Perfect score: 2287
 Sequence: 1 MELAALCRWGLLALLPFGA.....VGRGPDPPDAHVAVNLSRYEG 419

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 90 summaries

Database : PIR 78:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1878	82.1	1255	1 A24571	protein-tyrosine k
2	1608.5	70.3	1260	1 TVRTNU	protein-tyrosine k
3	1571	68.7	1254	2 I48161	p-185 precursor -
4	796.5	34.8	1223	1 TVCHLV	epidermal growth f
5	794	34.7	527	2 A42032	epidermal growth f
6	793	34.7	1210	1 GQHUE	epidermal growth f
7	789	34.5	1210	2 A53183	epidermal growth f
8	781.5	34.2	644	2 A35325	epidermal growth f
9	775	33.9	1308	2 A47253	epidermal growth f
10	735.5	32.2	1342	2 A36223	kinase-related tra
11	692	30.3	1339	2 JC4387	epidermal growth f
12	681.5	29.8	1166	1 S06142	protein-tyrosine k
13	575.5	25.2	843	2 A27131	epidermal growth f
14	419	18.3	1323	2 E88257	protein let-23 (im
15	419	18.3	1374	2 S70712	protein-tyrosine k
16	416	18.2	1369	2 S70713	protein-tyrosine k
17	415	18.1	1330	1 GQFE	epidermal growth f
18	363.5	15.9	366	2 D45558	epidermal growth f
19	363.5	15.9	1717	1 A45558	epidermal growth f
20	342.5	15.0	1363	2 T43220	insulin-like growth
21	331	14.5	333	2 B45558	epidermal growth f
22	331	14.5	342	2 C45558	epidermal growth f
23	294	12.9	2101	2 S57245	insulin receptor (
24	294	12.9	2148	1 A56081	insulin receptor
25	291	12.7	1477	2 T18534	protein-tyrosine k
26	278	12.2	1300	2 A36502	insulin receptor-k
27	272	11.9	540	2 B47417	insulin receptor-r
28	269.5	11.8	1607	2 T43212	insulin-like growth
29	263.5	11.5	1382	1 INUR	insulin receptor p

RESULT 1

A24571
 protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
 N/Alternate names: C-erb-B-2 protein precursor; Kinase-related transforming protein erb
 C/Species: Homo sapiens (man)
 C/Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text change 11-Jun-1999
 C/Accession: A24571; A25491; A44188; B44188; I59509; I57622

ALIGNMENTS

30	263.5	11.5	1383	2	A36080	insulin receptor p
31	261	11.4	1372	2	A34157	insulin receptor p
32	258	11.3	1390	2	T30346	insulin-like growth
33	257.5	11.3	1367	1	IGHUR1	insulin-like growth
34	252.5	11.0	1371	2	A33837	insulin-like growth
35	249	10.9	1268	2	B36502	insulin-like growth
36	235	10.3	329	2	A48805	insulin-like growth
37	226.5	9.9	183	2	JH0803	tyrosine kinase re
38	203	8.9	1846	2	T42047	insulin receptor h
39	160.5	7.0	1299	2	T43251	furin (EC 3.4.21.7
40	144	6.3	1548	2	S34583	serine proteinase
41	142	6.2	1959	1	AGRT	agrin - rat
42	136.5	6.0	1111	2	T26972	hypothetical prote
43	133.5	5.8	915	1	A48225	subtilisin-like pr
44	132	5.8	1574	2	T13954	MEGF6 protein - ra
45	131.5	5.7	915	2	B48225	probable protease
46	131.5	5.7	915	2	JC6148	subtilisin-like pr
47	129.5	5.7	417	2	T08724	hypothetical prote
48	129.5	5.7	899	2	G02426	subtilisin-like pr
49	129.5	5.7	1680	2	A43434	furin (EC 3.4.21.7
50	123.5	5.4	937	2	I53282	gene PACB4 protein
51	121.5	5.3	1737	2	T02029	MEGF8 protein - hu
52	120.5	5.3	294	2	T23682	hypothetical prote
53	120	5.2	837	2	S43656	furin (EC 3.4.21.7
54	120	5.2	942	2	D87803	protein bli-4D (im
55	120	5.2	2910	2	T42214	otogelin - mouse
56	119.5	5.2	667	2	A48579	trophozoite surfac
57	118.5	5.2	551	2	I46709	endothelial leukoc
58	118.5	5.2	932	2	I52527	PACE4A - mouse (fr
59	118.5	5.2	1620	2	T27283	hypothetical prote
60	118.5	5.2	3635	2	T10053	laminin alpha 5 ch
61	117.5	5.1	314	2	T27686	hypothetical prote
62	117.5	5.1	3623	2	T09456	intrinsic factor-B
63	117	5.1	469	1	S29126	properdin precursor
64	117	5.1	557	2	A48434	variant-specific s
65	117	5.1	3075	2	S14458	laminin alpha-1 ch
66	116.5	5.1	398	1	S24802	polyferredoxin 6x2
67	116.5	5.1	425	2	T18592	hypothetical prote
68	116.5	5.1	631	2	JC2345	kexin-like protein
69	116.5	5.1	969	1	A39490	subtilisin-like pr
70	116.5	5.1	975	2	JC5570	subtilisin-like pr
71	115	5.0	644	2	JC2346	kexin-like protein
72	115	5.0	1101	2	T16840	hypothetical prote
73	115	5.0	3106	1	S53868	laminin alpha-2 ch
74	114.5	5.0	482	2	JC5092	E-selectin - pig
75	114.5	5.0	600	2	T18593	hypothetical prote
76	114.5	5.0	1513	2	T23681	hypothetical prote
77	114	5.0	412	1	C37777	polyferredoxin 6x2
78	113.5	5.0	570	2	T37314	probable kexin (EC
79	113.5	5.0	1766	2	A42125	trophozoite cystei
80	113	4.9	782	2	A61625	tenascin-like prot
81	113	4.9	1122	2	I54237	protein-tyrosine k
82	113	4.9	1123	1	JN0712	protein-tyrosine k
83	113	4.9	1125	1	JH0771	protein-tyrosine k
84	113	4.9	1170	1	TSHUP1	thrombospondin 1 p
85	112.5	4.9	205	2	AG0360	probable dimethyl
86	112.5	4.9	513	2	D88991	protein apx-1 (imp
87	112.5	4.9	962	2	JC5571	subtilisin-like pr
88	111.5	4.9	572	2	T29880	hypothetical prote
89	111.5	4.9	1827	2	T34288	hypothetical prote
90	111	4.9	255	2	I38426	lymphocyte activat

R; Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.; Nature 319, 230-234, 1986
A>Title: Similarity of protein encoded by the human c-erbB-2 gene to epidermal growth factor
A;Reference number: A24571; MUID:86118663; PMID:3003577
A;Accession: A24571
A;Molecule type: mRNA
A;Residues: 1-1255 <YAM>
A;Cross-references: GB:X03363; NID:g31197; PIDN:CAA27060.1; PID:g31198
R; Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A>Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epidermal growth factor receptor gene
A;Reference number: A25491; MUID:86016729; PMID:2995967
A;Accession: A25491
A;Molecule type: DNA
A;Residues: 737-1031 <SEM>
A;Cross-references: GB:M11767; NID:g182163; PIDN:AAA35808.1; PID:g553282
R; Coussens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg, P.H.
Science 230, 1132-1139, 1985
A>Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosome 7 with the EGF receptor gene
A;Reference number: A44188; MUID:86070181; PMID:2999974
A;Accession: A44188
A;Molecule type: DNA
A;Residues: 740-910 <COU>
A;Cross-references: GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:g183989
A;Accession: B44188
A;Molecule type: mRNA
A;Residues: 1-517 <RALL>, 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A;Cross-references: GB:M11730; NID:g183986
R; King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 229, 974-976, 1985
A>Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A;Reference number: I59509; MUID:85272597; PMID:2992089
A;Accession: I59509
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 832-909 <REX>
A;Cross-references: GB:L29395; NID:g459807; PIDN:AAA35809.1; PID:g459808
R; Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A>Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptional regulation
A;Reference number: I57622; MUID:87286898; PMID:3039351
A;Accession: I57622
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-191 <TAL>
A;Cross-references: GB:M16792; NID:g183983; PIDN:AAA58637.1; PID:g553332
C;Comment: Amplification and overexpression of this erbB-related gene occurs in about 30% of human breast carcinomas
C;Genetics:
A;Gene: GDB:ERBB2; NGL; NEU; HER-2
A;Cross-references: GDB:120613; OMIM:164870
A;Map position: 17q21.1-17q21.1
A;Introns: 25/1; 75/3; 147/1; 883/3
A;Note: the list of introns is incomplete
C;Function:
A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C;Superfamily: epidermal growth factor receptor; protein kinase homology
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphatase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-1255/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>
F;22-653/Domain: extracellular #status predicted <EXT>
F;70-304/Domain: EGF receptor extracellular domain repeat <BE1>
F;395-605/Domain: EGF receptor extracellular domain repeat <BE2>
F;654-675/Domain: transmembrane #status predicted <TM>
F;676-1255/Domain: intracellular #status predicted <INT>
F;718-983/Domain: protein kinase homology <KIN>
F;726-734/Region: protein kinase ATP-binding motif
F;68,124,187,259,530,571,629/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F;753/Active site: Lys #status predicted
F;1139,1221,1222,1249/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 82.1%; Score 1878; DB 1; Length 1255;

Best Local Similarity 83.0%; Pred. No. 2.2e-120;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
QY 1 MEAALCRWGLLALPPGAASPTQVCTGDMKRLPASPETHLDMLRLHLYQGQVQGNL 60
DB 1 MEAALCRWGLLALPPGAASPTQVCTGDMKRLPASPETHLDMLRLHLYQGQVQGNL 60
QY 61 ELTYLPTNASLSPLQDIQVGVVLAHQVROVPLQRLRVKSTOLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSPLQDIQVGVVLAHQVROVPLQRLRVKSTOLFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTGASFGRLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQLA 180
DB 121 DPLNNTTPTVTGASFGRLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQLA 180
QY 181 LTLLDITNRSPACHPCSPCKGSCWGESSEDCOSLRTVCAGGCACCKPLPTDCHEQC 240
DB 181 LTLLDITNRSPACHPCSPCKGSCWGESSEDCOSLRTVCAGGCACCKPLPTDCHEQC 240
QY 241 AAGCTGPKESDCLACLFHFNHSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACP 300
DB 241 AAGCTGPKESDCLACLFHFNHSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACP 300
QY 301 YNYLSTDVSGSCTLVCPHNOEVTAEQTCCKSKPCAR-----CTHSLPPPAVFPV 355
DB 301 YNYLSTDVSGSCTLVCPHNOEVTAEQTCCKSKPCARVVCYGLGMEHLREVRVTSAN 360
QY 356 LRMPQG--PAHPVLSPLRPSWDLVSFYSLPLAPISPTSVPI-----SPVSVGRGPD 405
DB 361 IQEFACGCKIFGSLAPLPESFGDPAASNT--APLQPEQLQVFTLEETLITGVLYISAWPD 417
QY 406 --PDHVAVNLSRYEG 419
DB 418 SLPDLVSFQNLQVIRG 433
RESULT 2
TERTNU
Protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999
C;Accession: A24562; A61204
R; Bargmann, C.I.; Hung, M.C.; Weinberg, R.A.
Nature 319, 226-230, 1986
A;Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.
A;Reference number: A24562; MUID:86118662; PMID:3945311
A;Accession: A24562
A;Molecule type: mRNA
A;Residues: 1-1260 <BAR>
A;Cross-references: EMBL:X03362; NID:g56745; PIDN:CAA27059.1; PID:g56746
R; Masui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen
Carcinogenesis 12, 1975-1978, 1991
A;Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no 2-thiazolyl]formamide or N-methyl-N-nitrosourea.
A;Reference number: A61204; MUID:92035293; PMID:1682063
A;Accession: A61204
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 637-663 'V', 665-702 <MAS>
A;Note: authors translated the codon GCA for residue 25 as Val
C;Genetics:
A;Gene: neu
C;Superfamily: epidermal growth factor receptor; protein kinase homology
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphatase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>
F;658-680/Domain: transmembrane #status predicted <TM>
F;723-988/Domain: protein kinase homology <KIN>
F;731-739/Region: protein kinase ATP-binding motif
F;71,191,263,535,576,634/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;691/Binding site: phosphate (Thr) (covalent) #status predicted
F;758/Active site: Lys #status predicted
F;882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 70.3%; Score 1608.5; DB 1; Length 1260;
Best Local Similarity 85.0%; Pred No. 5.4e-102;
Matches 294; Conservative 15; Mismatches 36; Indels 1; Gaps 1;

Qy 1 MELALCRWGLLLALLPQGAASQVCTGDMKRLPASPETHLMLRHLVQCCVQVQNL 60
Db 4 MELAAWCRGFLTALLPQGIAGTQVCTGDMKRLPASPETHLMLRHLVQCCVQVQNL 63

Qy 61 ELTYLPTNASLFLQDIOEVQGYVLIANNQVQVPLQRLIRVGRQQLFEDNYALAVLNG 120
Db 64 ELTYYPANASLFLQDIOEVQGYVLIANNQVQVPLQRLIRVGRQQLFEDNYALAVLNG 123

Qy 121 DPLNNTTPTV-GASPGGLRELQRLSLEILKGGVLIQVNPOLCYQDTILWKDIFHNKQL 179
Db 124 DQDNVAASPTGTEGLRELQRLSLEILKGGVLIQVNPOLCYQDMVLKQVFRKNQL 183

Qy 180 ALTLIDTNRSRACHPCSPCKGSRGWSSEDCQSITRTVCAGGCKGRLPTDCCHQC 239
Db 184 AFVDITNRSRACPPCAPACDNHCWSPEDCQILGTICTSGCARCKGRLPTDCCHQC 243

Qy 240 CAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTAC 299
Db 244 CAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMHNPEGRYTFGASCVTTC 303

Qy 300 PNYLSTVGSCTLLVCPLEHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTSL 345
Db 304 PNYLSTVGSCTLLVCPLEHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTSL 349

RESULT 3
148161
P-185 precursor - golden hamster
C:Species: Mesocricetus auratus (golden hamster)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
C:Accession: I48161
R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishikawa, Gene 140, 251-255, 1994
A>Title: Cloning and activation of the Syrian hamster neu proto-oncogene.
A:Reference number: 148161; PMID:94193007; PMID:7908275
A:Accession: I48161
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1254 <RSS>
A:Cross-references: GB:D16295; NID:g493236; PIDN:5MA03801.1; PID:g747595
C:Genetics:
A:Gene: neu
A:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP
F:718-983/Domain: protein kinase homology <KIN>
F:726-734/Region: protein kinase ATP-binding motif

Query Match 68.7%; Score 1571; DB 2; Length 1254;
Best Local Similarity 74.4%; Pred. No. 1.9e-99;
Matches 299; Conservative 26; Mismatches 67; Indels 10; Gaps 3;

Qy 1 MELALCRWGLLLALLPQGAASQVCTGDMKRLPASPETHLMLRHLVQCCVQVQNL 60
Db 1 MELAAWCRGFLTALLPQGIAGTQVCTGDMKRLPASPETHLMLRHLVQCCVQVQNL 60

Qy 61 ELTYLPTNASLFLQDIOEVQGYVLIANNQVQVPLQRLIRVGRQQLFEDNYALAVLNG 120
Db 64 ELTYYPANASLFLQDIOEVQGYVLIANNQVQVPLQRLIRVGRQQLFEDNYALAVLNG 120

Qy 121 DPLNNTTPTV-GASPGGLRELQRLSLEILKGGVLIQVNPOLCYQDTILWKDIFHNKQL 180
Db 124 DQDNVAASPTGTEGLRELQRLSLEILKGGVLIQVNPOLCYQDMVLKQVFRKNQL 180

Qy 181 LTLIDTNRSRACHPCSPCKGSRGWSSEDCQSITRTVCAGGCKGRLPTDCCHQC 240
Db 184 AFVDITNRSRACPPCAPACDNHCWSPEDCQILGTICTSGCARCKGRLPTDCCHQC 240

Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTAC 300

241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTTC 300
Qy 301 YNYLSTVGSCTLLVCPLEHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTTC 355
Db 301 YNYLSTVGSCTLLVCPLEHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTTC 360
Qy 356 LRMQPG--PAHPVLSFLRPSMDLVSAFYSLPLAPLSPTSPVI 395
Db 361 IQEAGCKKIFGSLAFLPESD---GNPSSGIALPLTPEQLQV 399

RESULT 4
TVCHLV
epidermal growth factor receptor precursor - chicken
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB
C:Species: Gallus gallus (chicken)
C>Date: 28-Feb-1986 #sequence_revision 05-May-1995 #text_change 04-Feb-2000
C:Accession: A27720; A00643
R:Lax, I.; Johnson, A.; Howk, R.; Sap, J.; Bellot, F.; Winkler, M.; Ullrich, A.; Vennet Mol. Cell. Biol. 8, 1970-1978, 1988
A>Title: Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in m A:Reference number: A27720; PMID:88261272; PMID:3260329
A:Accession: A27720
A:Molecule type: mRNA
A:Residues: 1-1223 <LAX>
A:Cross-references: GB:M20386
R:Nilsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Raines, Cell 41, 719-726, 1985
A>Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and p A:Reference number: A00643; PMID:85228222; PMID:2988784
A:Accession: A00643
A:Molecule type: mRNA
A:Residues: 585-1223 <NIL>
A:Cross-references: GB:M10066
C:Genetics:
A:Gene: erbB
A:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor; specific protein kinase
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-1223/Product: epidermal growth factor receptor #status predicted <MAT>
F:31-654/Domain: extracellular #status predicted <EXT>
F:81-307/Domain: EGF receptor extracellular domain repeat <BE1>
F:397-610/Domain: EGF receptor extracellular domain repeat <BE2>
F:655-677/Domain: transmembrane #status predicted <TM>
F:678-1223/Domain: intracellular #status predicted <INT>
F:719-984/Domain: protein kinase homology <KIN>
F:727-735/Region: protein kinase ATP-binding motif
F:136,202,280,361,376,422,575,580,615,635/Binding site: carbohydrate (Thr) (covalent) # F:192,650/Binding site: carbohydrate (Ser) (covalent) #status predicted
F:687/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F:754/Active site: Lys #status predicted
F:1100,1183,1208/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #sta

Query Match 34.8%; Score 796.5; DB 1; Length 1223;
Best Local Similarity 45.5%; Pred. No. 1.3e-46;
Matches 155; Conservative 55; Mismatches 112; Indels 19; Gaps 7;

Qy 8 RWGLLEALLPGAR-----STQVCTGDMKRLPASPETHLMLRHLVQCCVQVQNL 61
Db 13 RGAALVLLLLGVALCSAEBKVCQGTNNKLTQGHVEHFTSLQRYNCEVLSNLE 72

Qy 62 LTYLPTNASLFLQDIOEVQGYVLIANNQVQVPLQRLIRVGRQQLFEDNYALAVLNGD 121
Db 73 ITVYHNRDLTFLKTIQEVAGYVLIANNQVQVPLQRLIRVGRQQLFEDNYALAVLSNVH 132

Qy 122 DPLNNTTPTV-GASPGGLRELQRLSLEILKGGVLIQVNPOLCYQDTILWKDIFHNKQL 181
Db 124 DQDNVAASPTGTEGLRELQRLSLEILKGGVLIQVNPOLCYQDMVLKQVFRKNQL 180

Qy 181 LTLIDTNRSRACHPCSPCKGSRGWSSEDCQSITRTVCAGGCKGRLPTDCCHQC 240
Db 184 AFVDITNRSRACPPCAPACDNHCWSPEDCQILGTICTSGCARCKGRLPTDCCHQC 240

Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTAC 300

```

Db      183  TVLDFASNLSSCPKHPNCTEDHCHWGAGEQNCQTLTVKVIQAQCQSGRCGRKVPSPDCCHNQ 243
QY      240  CAAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTTFESMPNPEGRYTFGASCVTAC 299
Db      243  CAAGCTGPPSDCLACRFRDDATCKDTCPLVLVNTTYQMDVNPBGKYSFGATCVREC 302
QY      300  PNYLSTDVGSCTLVCPLENQEVTAEDTQRCCKSPCAR 340
Db      303  PHNYVVYTDHSGSVRSNTDTYEV-EENGVRKCKKCDGLCSK 342

RESULT 5
A42032
epidermal growth factor receptor - chicken
C:Species: Gallus gallus (Chicken)
C:date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 18-Jun-1999
C:Accession: A42032
R:Flickinger, T.W.; Malth, N.J.; Kung, H.J.
Mol. Cell. Biol. 12, 883-893, 1992
A:title: An alternatively processed mRNA from the avian c-erbB gene encodes a
A:Reference number: A42032; MUID:92123214; PMID:1732751
A:Accession: A42032
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-527 rFLI>
A:Cross-references: GB:M7637; MID:9211737; PIDN:AAA48759.1; PID:G9211738
A:Experimental source: liver
A>Note: sequence extracted from NCBI backbone (NCBIN:76892, NCBIPI:76893)
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; growth factor receptor

Query Match 34.7%; Score 794; DB 2; Length 527;
Best Local Similarity 46.0%; Pred. No. 8e-47;
Matches 155; Conservative 55; Mismatches 109; Indels 18; Gaps 7

QY      11  LLLALLPGCAAST- ---QVCTGTDMLKLPASPETHLDMLRHLYQCGVQVGNLSTLYL 65
Db      20  LLLLLLGRVALCSAVEEKVYCGQTNKLTQLGHVEDFTSLQRYNANCEVVLNLAITY 79

QY      66  PTNASLFLQDIOEYGVYLIHAHQVQVPIQLRIVRGTQLPEDNYALAVLQNGOPLNN 125
Db      80  EHNRDLTFTKTIQEVAGYVLIALNMVDVPILENLQIIRGNVLYDNSPALAVLSNYH-NKK 138

QY      126  TTPVTGASPGGLREQLSLTEILKGGVLIQRNPOLCYODITLWKDILFHKNKQALATLID 185
Db      139  TQ-----GLRELPMKLSILNGGVKISNNPKLQMDTVLWMDLIDITSRK-PLTVLD 189

QY      186  -TNRSPACHPCSPMKGRCWGESSEDCOSITRTVCAAGCA-RCKGPLPTDCCHQCAAG 243
Db      190  FASNLSSCPKHPNCTEDHCHWGAGEQNCQTLTVKVICAQCQSGRCGRKVPSPDCCHQCAAG 249

QY      244  CTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTTFESMPNPEGRYTFGASCVTACPNY 303
Db      250  CTGPRESDCLACRFRDDATCKDTCPLVLVNTTYQMDVNPBGKYSFGATCVREPHNY 309

QY      304  LSTDVGSCTLVCPLENQEVTAEDGTQRCCKSPCAR 340
Db      310  VYTDHSGSVRSNTDTYEV-EENGVRKCKKCDGLCSK 345

```

A:Molecule type: mRNA
A:Residues: 1-1210 <ULL>
A:Cross-references: EMBL:X00588; NID:G31113; PIDN:CAA25240.1; PID:G757924
A:Note: The authors translated the codon AAG for residue 540 as Asn
R:Ishii, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.
Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985
A:Title: Characterization and sequence of the promoter region of the human epidermal
A:Reference number: A25772; MUID:85270438; PMID:2991899
A:Accession: A25772
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-29 <ISH>
A:Cross-references: GB:M11234; NID:G181981; PIDN:AAA52370.1; PID:G553272
R:Haley, J.; Whittle, N.; Bennett, P.; Kinchington, D.; Ullrich, A.; Waterfield, M.
Oncogene Res. 1, 375-396, 1987
A:Title: The human EGF receptor gene: structure of the 110 kb locus and identification
A:Reference number: S30024; MUID:88217333; PMID:3329716
A:Accession: S30024
A:Molecule type: DNA
A:Residues: 1-29 <HA2>
A:Cross-references: EMBL:X06370; NID:G31118; PIDN:CAA29668.1; PID:G31119
R:Haley, J.D.; Waterfield, M.D.
J. Biol. Chem. 266, 1746-1753, 1991
A:Title: Contributory effects of de Novo transcription and premature transcript termi
A:Reference number: A38672; MUID:91107677; PMID:1988448
A:Accession: A38672
A:Molecule type: DNA
A:Residues: 1-29 <HA>
A:Cross-references: GB:M384425; NID:G181977; PIDN:AAA63171.1; PID:G553271
A:Experimental source: carcinoma cell line A431-7
R:Xu, Y.; Ishii, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.; M
Nature 309, 806-810, 1984
A:Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RN
A:Reference number: A00642; MUID:84245835; PMID:6330563
A:Accession: A00642
A:Molecule type: mRNA
A:Residues: 'RCAMRA', 150-187, 'KSVTQAV', 195, 'M', 197, 'A', 199-222, 'S', 224-304, 'RA', 307-3
, '798-799', 'TD', '802-811', 'B', '813-942 <XUY>
A:Experimental source: A431 human carcinoma cells, which have large numbers of EGF rec
R:Lin, C.R.; Chen, W.S.; Krueger, W.; Stolarsky, L.S.; Weber, W.; Evans, R.M.; Verma,
Science 224, 843-848, 1984
A:Title: Expression cloning of human EGF receptor complementary DNA: gene amplificatio
A:Reference number: A43615; MUID:84196372; PMID:6326261
A:Accession: A43615
A:Molecule type: mRNA
A:Residues: 713-964 <LIN>
A:Experimental source: epidermoid carcinoma cell line A431
R:Simmen, F.A.; Gope, M.L.; Schulz, T.Z.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.
Biochem. Biophys. Res. Commun. 124, 125-132, 1984
A:Reference number: A23062; MUID:85046483; PMID:6093780
A:Accession: A23062
A:Molecule type: mRNA
A:Residues: 1028-1210 <STM>
R:Weber, W.; Gill, G.N.; Speiss, J.
Science 224, 294-297, 1984
A:Reference number: A05281; MUID:84172183; PMID:6324343
A:Accession: A05281
A:Molecule type: protein
A:Residues: 25-30, 'S', 32-51, 454-467 <WEB>
R:Russo, M.W.; Lukas, T.J.; Cohen, S.; Staros, J.V.
J. Biol. Chem. 260, 5205-5208, 1985
A:Title: Identification of residues in the nucleotide binding site of the epidermal gr
A:Reference number: A60143; MUID:85182650; PMID:2985580
A:Accession: A60143
A:Molecule type: protein
A:Residues: 740-744, 'X', 746-747 <RUS>
R:Mróczkowski, B.; Moesig, G.; Cohen, S.
Nature 309, 270-273, 1984
A:Title: ATP-stimulated interaction between epidermal growth factor receptor and super
A:Reference number: A38023; MUID:84191554; PMID:6325948
A:Contents: annotation; receptor activity
A:Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA
R:Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C.

Cell 59, 33-43, 1989
A:Title: Functional independence of the epidermal growth factor receptor from a domain 2
A:Reference number: A4331; MUID:90003233; PMID:2790960
A:Contents: annotation; internalization signal
A:Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor
C:Genetics:
A:Gene: GDB:EGFR
A:Cross-references: GDB:120610; OMIM:131550
A:Map position: 7p12.3-7p12.1
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-1210/Product: EGF receptor #status predicted <EXT>
F:25-645/Domain: extracellular #status predicted <EXT>
F:25-300/Domain: EGF receptor extracellular domain repeat <EE1>
F:390-600/Domain: EGF receptor extracellular domain repeat <EE2>
F:646-668/Domain: transmembrane #status predicted <TM>
F:669-1210/Domain: intracellular #status predicted <INT>
F:710-975/Domain: protein kinase homology <KIN>
F:718-726/Region: protein kinase ATP-binding motif
F:999-1046/Region: coated-pit mediated internalization signal
F:1047-1210/Region: inhibitory
F:128,175,352,413,444,528,603/Binding site: carbohydrate (Asn) (covalent) #status predic
F:745/Active site: Lys #status experimental

Query Match 34.7%; Score 793; DB 1; Length 1210;
Best Local Similarity 45.3%; Pred. No. 2.2e-46;
Matches 151; Conservative 50; Mismatches 118; Indels 14; Gaps 4;

Qy 11 LLLALLPGAA--STOVCTGDMKLRLPASPETHLMDRLHYQGVQVQGNLEITYPTN 68
Db 14 LLAALCPASRALEKKVCGQTSNRLTQLGTFEDHFLSLQRMNCEVVLGNLEITYVORN 73
Qy 69 ASLSFLQDIQEVQGVYLIHNRQVPLQRLIRVRGTQLFEDNYALVLDNGDPLNNTTP 128
Db 74 YDLSEFLTKIQEVAGYVLIHNRQVPLQRLIRVRGTQLFEDNYALVLDNGDPLNNTTP 126
Qy 129 VTGASPGGLRELQRLSLEILKGGVLIQNRNPOLCYQDTILWKDIHKNQALTLITNR 188
Db 127 ---ANKTKELPRLNLEIILHGAVRFSNNPACNVESIQWRDIVSSDFLSNMDFQNH 183
Qy 189 SRACHPCPMCKGRCSWESSEDCQSLRTRVCAGCA-RCKGPLETDCCHSOCAAGCTGP 247
Db 184 LGSQKCDPSPNCSWAGAGENCQKTIICAQCSGRCKGKSPDCCHQCAAGCTGP 243
Qy 248 KHSQCLACLFHNSGICELCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLTSD 307
Db 244 RESDCLVCRKFRDEATCKDTCPPLMLYNPTTYQMDVNPPEGKYSFGATCVKXKCPNYYVTD 303
Qy 308 VGSCTLVCLPHNQEVTAEDGTQRCCKSPCAR 340
Db 304 HGSCVRACGADSYEM-BEDGVRKCKCKEGPCRK 335

RESULT 7
A53183
epidermal growth factor receptor precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text change 18-Jun-1999
C:Accession: A53183; A43818; S24942; A28941; S45325; I49643
R:Luettke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A.;
Genes Dev. 8, 399-413, 1994
A:Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor
A:Reference number: A53183; MUID:94170986; PMID:8125255
A:Accession: A53183
A:Molecule type: mRNA
A:Residues: 1-1210 <LUE>
A:Cross-references: GB:U03425
R:Avivi, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Givol, D.; Morse, B.
Oncogene 6, 673-676, 1991
A:Title: Comparison of EGF receptor sequences as a guide to study the ligand binding sit
A:Reference number: A43818; MUID:91232866; PMID:2030916
A:Accession: A43818

A:Molecule type: mRNA
A:Residues: 1-714 <AVI>
A:Cross-references: GB:X59698
R:Eslinger, D.P.; Serrero, G.
submitted to the EMBL Data Library, June 1992
A:Reference number: S24942
A:Accession: S24942
A:Molecule type: mRNA
A:Residues: 989-971, 'K', 973-1115, 'D' <EIS>
A:Cross-references: EMBL:Z12608
R:Heisermann, G.J.; Gill, G.N.
J. Biol. Chem. 263, 13152-13158, 1988
A:Title: Epidermal growth factor receptor threonine and serine residues phosphorylated
A:Reference number: A28941; MUID:88330814; PMID:3138233
A:Accession: A28941
A:Molecule type: protein
A:Residues: 689-694, 'X', 696-704, 'L', 706-707, 989-992, 'XX', 995-996, 'X', 998-1000, 1002-1003;
R:Hibbs, M.L.; Dunn, A.R.; Alexander, W.S.
submitted to the EMBL Data Library, April 1994
A:Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Receptor
A:Reference number: S45325
A:Accession: S45325
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-971, 'K', 973-1210 <VER>
A:Cross-references: EMBL:X78987; NID:G488830; PIDN:CAA55587.1; PID:G488831
R:Paria, B.C.; Das, S.K.; Andrews, G.K.; Dey, S.K.
Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993
A:Title: Expression of the epidermal growth factor receptor gene is regulated in mouse
A:Reference number: I49643; MUID:93126380; PMID:7678348
A:Accession: I49643
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 12-20, 22-132 <RES>
A:Cross-references: GB:L06864; NID:G193001; PIDN:AAA53029.1; PID:G567201
C:Genetics:
A:Gene: EGFR
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; growth factor receptor; kinase-related transforming protein; phospho
F:1-24/Domain: signal sequence #status predicted <SIG>
F:648-670/Domain: transmembrane #status predicted <TM>
F:712-977/Domain: protein kinase homology <KIN>
F:720-728/Region: protein kinase ATP-binding motif
F:680,695/Binding site: phosphate (Thr) (covalent) #status experimental
F:697,1070,1071/Binding site: phosphate (Ser) (covalent) #status experimental
F:993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental
F:1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental
F:1197/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 34.5%; Score 789; DB 2; Length 1210;
Best Local Similarity 46.3%; Pred. No. 4.2e-46;
Matches 156; Conservative 43; Mismatches 116; Indels 22; Gaps 6;

Qy 11 LLLALLPGAA--STOVCTGDMKLRLPASPETHLMDRLHYQGVQVQGNLEITYPTN 68
Db 14 LLAALCAAGALEKKVCGQTSNRLTQLGTFEDHFLSLQRMNCEVVLGNLEITYVORN 73
Qy 69 ASLSFLQDIQEVQGVYLIHNRQVPLQRLIRVRGTQLFEDNYALVLDNGDPLNNTTP 128
Db 74 YDLSEFLTKIQEVAGYVLIHNRQVPLQRLIRVRGTQLFEDNYALVLDNGDPLNNTTP 124
Qy 129 VTGASPGGLRELQRLSLEILKGGVLIQNRNPOLCYQDTILWKDIHKNQALTLITLI 184
Db 125 -YGNRTGLRELPMRLNLEIILHGAVRFSNNPACNVESIQWRDIVSSDFLSNMDFQNH 180
Qy 185 DTRSRACHPCPMCKGRCSWESSEDCQSLRTRVCAGCA-RCKGPLETDCCHSOCAAG 243
Db 181 -QSHPSCKPDCSPNCSWAGAGENCQKTIICAQCSGRCKGKSPDCCHQCAAG 239
Qy 244 CTGPKHSQCLACLFHNSGICELCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNY 303
Db 240 CTGPRSDCLVCRKFRDEATCKDTCPPLMLYNPTTYQMDVNPPEGKYSFGATCVKXKCPN 299

```
Qy 304 LSTDVGSCTLVCPLNHNOEVTAEEDGTQRCCKSPCAR 340
      |||||
Db 300 VTDHGSVCVRAGDPDYEV-EEDGIRKCKCDGCPCK 335

RESULT 8
A36325
epidermal growth factor receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Jan-1991 #sequence_revision 25-Jan-1991 #text_change 10-Oct-1997
C:Accession: A36325
R:Petch, L.A.; Harris, J.; Raymond, V.W.; Blasband, A.; Lee, D.C.; Earp, H.S.
Mol. Cell. Biol. 10, 2973-2982, 1990
A:Title: A truncated, secreted form of the epidermal growth factor receptor is encoded by a complementary DNA clone from rat epidermis
A:Reference number: A36325; MUID:9025888; PMID:2342466
A:Accession: A36325
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-644 <P>
C:Cross-references: GB:W37394
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: alternative splicing; ATP; growth factor receptor

Query Match 34.2%; Score 781.5; DB 2; Length 644;
Best Local Similarity 46.0%; Pred. No. 7e-46;
Matches 157; Conservative 44; Mismatches 117; Indels 23; Gaps 7;

Qy 3 LAALCRWGLLLALLPPGA-ASTQVCTGDMKRLPASPEHLDMLRHLYOGQVQGNLE 61
      |||||
Db 15 LAALCAAG-----GALEBKVKCGTSSNKLTLQGFEDHFUSLQRMFNCEVJGNLE 66

Qy 62 LTYLPTNASLFLODIQEVQGVLIHNOVQVPLQRLIRIVRGTPQLFEDNYALAVLDNGD 121
      |||||
Db 67 ITYVQRYVDISFLAKTIQEVACVYLIALTVERIPLENLIQINGNALYENTYALAVLSN-- 124

Qy 122 PLNNTPTVGTASPGLELQRLSLTEILKGVLIQRLNPOCYQDTILWKDIFHKNNQAL 181
      |||||
Db 125 -----YGNKTKGLELPMNQLQELIGAVRFSNPILCMNMTIQRDIV-QDVFLSN 175

Qy 182 TLIDNRS-RACHCSPMCKSGRCWGESSEDCQSLTRTVACAGCA-RCKGPLEPDCCHEQ 239
      |||||
Db 176 MSMDVQRHLTKPCKDPSCPNCSKWRGEENCQKLTICAAQCSCRGRSPSCCHNQ 235

Qy 240 CAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTAC 299
      |||||
Db 236 CAAGCTGERSDCLVCHFRDEATCKTCCPELMLYNPTTYQMDVNPGRKYSFGATCVKCK 295

Qy 300 PNYLSTDVGSCTLVCPLNHNOEVTAEEDGTQRCCKSPCAR 340
      |||||
Db 296 PRNYVTDHGSVCVRAGDPDYEV-EEDGVSKCKCDGCPCK 335

RESULT 9
A47253
epidermal growth factor receptor, HER4 - human
C:Species: Homo sapiens (man)
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C:Accession: A47253
R:Plowman, G.D.; Culouscou, J.M.; Whitney, G.S.; Green, J.M.; Carlton, G.W.; Foy, L.; Ne
Proc. Natl. Acad. Sci. U.S.A. 90, 1746-1750, 1993
A:Title: Ligand-specific activation of HER4/p180erbB4, a fourth member of the epidermal
A:Reference number: A47253; MUID:93189574; PMID:8383326
A:Accession: A47253
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-1308 <P>
C:Cross-references: GB:L07868; NID:g337359; PIDN:AA859446.1; PID:g337360
A>Note: sequence extracted from NCBI backbone (NCBIP:126842)
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; growth factor receptor
F:716-981/Domain: protein kinase homology <KIN>
F:724-732/Region: protein kinase ATP-binding motif
```

```
Query Match 33.9%; Score 775; DB 2; Length 1308;
Best Local Similarity 45.1%; Pred. No. 4.1e-45;
Matches 157; Conservative 45; Mismatches 124; Indels 22; Gaps 7;

Qy 9 WGLLLALLPPGAA---STQVCTGTDMLRLPASPEHLDMLRHLYOGQVQGNLELYT 64
      |||||
Db 8 WTVVSLVVAAGTVQPSDSQSCAGTENKLSLSLEQQVYRALRYKYYENCEVWGNLEITS 67

Qy 65 LPTNASLFLODIQEVQGVLIHNOVQVPLQRLIRIVRGTPQLFEDNYALAVLDNGDPLN 124
      |||||
Db 68 IEHNRDLSPRSRVREVTGYLVALNQFRYLPLENRIITGTLYEDRYALALFLNRYKDG 127

Qy 125 NTPPTVGTASPGLELQRLSLTEILKGVLIQRLNPOCYQDTILWKDIFHKNNQALTLI 184
      |||||
Db 128 NF-----GLQELGLKMLTEILNGVYVDQNKELCYADIIHQDIVENPWSNLTIV 178

Qy 185 DYNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVACAGC-ARCKGPLETDCCHQCAAG 243
      |||||
Db 179 STNGSGGCRCHKSCGTG-RCWGPTEHNCOTLRTVCAEQDCRCYGPVYSDCCHRECAAG 237

Qy 244 CTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNY 303
      |||||
Db 238 CSGPKDTCFACNFMDSGACTVQCPQTFVYNTPTFQLEHNFNAXKTYGAFCKKCPHF 297

Qy 304 LSTDVGSCTLVCPLNHNOEVTAEEDGTQRCCKSPCAR-----GTHSL 346
      |||||
Db 298 V-VDSSSVCVRACPSRRMEV-EENGKMKCKPCTDICKACDGGTGTGSLM 343

RESULT 10
A36223
kinase-related transforming protein (erbB3) (BC 2.7.1.1-) precursor - human
C:Species: Homo sapiens (man)
C:Date: 04-Oct-1991 #sequence_revision 13-Jan-1993 #text_change 17-Nov-2000
C:Accession: A36223; I59164
R:Kraus, M.H.; Issing, W.; Miki, T.; Popescu, N.C.; Aaronson, S.A.
Proc. Natl. Acad. Sci. U.S.A. 86, 9193-9197, 1989
A:Title: Isolation and characterization of ERBB3, a third member of the ERBB/epidermal
A:Reference number: A36223; MUID:90083234; PMID:2687875
A:Accession: A36223
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1342 <K>
A:Cross-references: GB:M29366
R:Plowman, G.D.; Whitney, G.S.; Neubauer, M.G.; Green, J.M.; McDonald, V.L.; Todaro, G
Proc. Natl. Acad. Sci. U.S.A. 87, 4905-4909, 1990
A:Title: Molecular cloning and expression of another epidermal growth factor receptor-
A:Reference number: I59164; MUID:90311312; PMID:2164210
A:Accession: I59164
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-559, 'G', 561-957, 'F', 959-1063, 'G', 1065-1342 <RES>
A:Cross-references: GB:M34309; NID:g183990; PIDN:AAA35979.1; PID:g306841
C:Genetics:
A:Gene: GDB:ERBB3; HER3
A:Cross-references: GDB:119880; OMIM:190151
A:Map position: 12q13-12q13
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homoc
C:Keywords: ATP; phosphotransferase
F:707-972/Domain: protein kinase homology <KIN>
F:715-723/Region: protein kinase ATP-binding motif

Query Match 32.2%; Score 735.5; DB 2; Length 1342;
Best Local Similarity 44.0%; Pred. No. 2.1e-42;
Matches 150; Conservative 49; Mismatches 121; Indels 21; Gaps 9;

Qy 10 GLLALLPPGAA---STQVCTGTDMLRLPASPEHLDMLRHLYOGQVQGNLELYLPT 67
      |||||
Db 11 GLIFSLARGSEVNSQAVCPGTGLNGLSVTGDENQYCTLYKLYERCEVWGNLEIVLTGH 70

Qy 68 NASLSFLQDIQEVQGVLIHNOVQVPLQRLIRIVRGTPQLFEDNYALAVLDNGDPLNNT 127
      |||||
Db 71 NADLSFLQIREVTGYLVVAMNEFSTLPLNLAIVRGTVQYDCKFAIFW-----LNYNT 125
```


QY 128 PVTGASPGGLRLQLRSITLTKGVLIQRNPOLCYQDTILWKDIFPHKNNQLALTIDTN 187
 DB 126 ----NSSHALRQLRLTQLTSLGGVYIEKNDKLCMDITDWRDVRD---AEIVVKD 178
 QY 188 RSRACHPCSPCKGRGSESSDCQSLRTVTCAGGC-ARCKGRLPTDCCHEQCAAGCTG 246
 DB 179 NGRSPPCHEVCKG-RCWPGSEDCQTLTKTICAPQCHGCHGVPNQCCHDECAAGCGG 237
 QY 247 FKHSCLACLHFNHSGICELHCPALVYNTDFTFESMPNPGRYTFGASCVTACPNYILST 306
 DB 238 PQDTCFACRHFNDGACVPCRPQPLVYNKLTFLQEPNPHTKYQYGGVGVASCPHFV-V 296
 QY 307 DVGSTVLCPLHNOEVTADGTQCEKSKPCAR---GTHS 344
 DB 297 DQTSVCRAPPDKMEVD-KHGLKMCPECGGLCPKACEGTGS 336

RESULT 11
 Jc4387
 epidermal growth factor receptor homolog precursor - rat
 N:Alternate names: ErbB3 protein; HER3 protein
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 17-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 13-Nov-1998
 C:Accession: Jc4387
 R:Hellwyer, N.J.; Kim, H.H.; Greaves, C.H.; Sierke, S.L.; Koland, J.G.
 Gene 165, 279-284, 1995
 A:Title: Cloning of the rat ErbB3 cDNA and characterization of the recombinant protein.
 A:Reference number: Jc4387; MUID:96036535; PMID:8522190
 A:Accession: Jc4387
 A:Molecule type: mRNA
 A:Residues: 1-1339 <HEL>
 A:Cross-references: GB:U29339; NID:9915389; PID:9915390
 A:Experimental source: liver
 A:Note: The authors translated the codon AAC for residue 369 as Thr and GTT for residue 454.
 C:Comment: This protein is a functional heregulin receptor that transduces signals to the cell.
 C:Genetics:
 A:Gene: ErbB3
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
 C:Keywords: ATP; growth factor receptor; liver; phosphoprotein; transmembrane protein
 F:1-19/Domain: signal sequence status predicted <SIG>
 F:20-1339/Product: epidermal growth factor homolog #status predicted <MAT>
 F:640-653/Domain: transmembrane #status predicted <TM>
 F:640-653/Domain: protein kinase homolog <KIN>
 F:705-970/Domain: protein kinase ATP-binding motif
 F:713-721/Region: protein kinase ATP-binding motif
 F:939,1051,1156,1194,1196,1219,1257,1259,1273,1286,1325/Binding site: phosphate (Tyr) (c

Query Match 30.3%; Score 692; DB 2; Length 1339;
 Best Local Similarity 41.8%; Pred. No. 1.9e-29;
 Matches 146; Conservative 46; Mismatches 131; Indels 26; Gaps 10;

QY 3 LAALCRWGLLLALLPFGAA---STQVCTGTDMLRLPASPEHLDMLRHLHYQGVQVQGN 59
 DB 7 LQVLC---FLLSLARGSEMGNSQAVCPQTLNGLSVTGDADNQYQTLVLYKECEVVMGN 62
 QY 60 LELTYLPTNASLSFLQDIOEYGVYLIQHNQVQVPLQRLIRVGRGTQFLFDNYALAVLDN 119
 DB 63 LEIVLTGHNADLSFLQIRVETAYVLIQHNQVQVPLQRLIRVGRGTQFLFDNYALAVLDN 120
 QY 120 GDPNNITPTVTCASPGGLRLQLRSITLTKGVLIQRNPOLCYQDTILWKDIFPHKNNQL 179
 DB 121 ----LNYNT---NSSHALRQLKFTQLTSLGGVYIEKNDKLCMDITDWRDVRV--- 170
 QY 180 ALTLIDTNRSRACHPCSPCKGRGSESSDCQSLRTVTCAGGC-ARCKGRLPTDCCHE 238
 DB 171 GAELIVKNGANGCPPCHEYCKG-RCWPGPDDCQTLTKTICAPQCHGCHGVPNQCCHD 229
 QY 239 QCAAGCTGPKHSDCLACLHFNHSGICELHCPALVYNTDFTFESMPNPGRYTFGASCVTA 298
 DB 230 ECAGCGSGPQDIDCFACRFNDSGACVPCRPPLVYNKLTFLQEPNPHTKYQYGGVGVAS 289
 QY 299 CPNYLSTDVGSCTVLCPLHNOEVTADGTQCEKSKPCAR---GTHS 344

DB 290 CPHNFV-VDQTFVCRAPPDKMEVD-KHGLKMCPECGGLCPKACEGTGS 336

RESULT 12 S06142

protein-tyrosine kinase (EC 2.7.1.112) mrk-Y precursor - southern platyfish
 N:Alternate names: epidermal growth factor receptor homolog; kinase-related transform
 C:Species: Xiphophorus maculatus (southern platyfish)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 18-Feb-2000
 C:Accession: S06142; S13809
 R:Wittbrodt, J.; Adam, D.; Malitschek, B.; Maeueller, W.; Raulf, F.; Telling, A.; Rober
 Nature 341, 415-421, 1989
 A:Title: Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu 1
 A:Reference number: S06142; MUID:90015140; PMID:2797166
 A:Accession: S06142
 A:Molecule type: DNA
 A:Residues: 1-1166 <MIT>
 A:Cross-references: EMBL:X16891; NID:965290; PIDN:CAA34770.1; PID:965291
 R:Adam, D.; Maeueller, W.; Scharf, M.
 Oncogene 6, 73-80, 1991
 A:Title: Transcriptional activation of the melanoma inducing xmrk oncogene in Xiphopho
 A:Reference number: S13807; MUID:91125882; PMID:1846957
 A:Accession: S13809
 A>Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 821-1025, 'N', 1027-1098, 'A', 1100-1166 <ADA>
 A:Cross-references: EMBL:X56319; NID:965284; PIDN:CAA39763.1; PID:965285
 C:Genetics:
 A:Gene: mrk
 A:Map position: Y
 A:Introns: 872/3; 898/1; 947/1; 979/3; 1025/3; 1056/1
 C:Superfamily: epidermal growth factor receptor; protein kinase homolog
 C:Keywords: ATP; growth factor receptor; phosphotransferase; transmembrane protein; ty
 F:1-25/Domain: signal sequence #status predicted <SIG>
 F:26-1166/Product: kinase-related transforming protein (Tu) #status predicted <MAT>
 F:707-972/Domain: protein kinase homolog <KIN>
 F:713-723/Region: protein kinase ATP-binding motif

Query Match 29.8%; Score 681.5; DB 1; Length 1166;
 Best Local Similarity 42.4%; Pred. No. 8.7e-39;
 Matches 145; Conservative 46; Mismatches 136; Indels 15; Gaps 8;

QY 4 AALCRWGLLLALLPFGAAST---OVCTGTDMLRLPASPEHLDMLRHLHYQGVQVQGN 59

DB 8 AALIQ---LLLVLSISRCSTDPDRKVCQGTSTNQMT---LDNHLYKMKMYSGCNVLEN 62

QY 60 LELTYLPTNASLSFLQDIOEYGVYLIQHNQVQVPLQRLIRVGRGTQFLFDNYALAVLDN 119

DB 63 LEITYTQENQDLSFLQSIQEVGVYLIQHNQVQVPLQRLIRVGRGTQFLFDNYALAVLDN 122

QY 120 GDPNNITPTVTCASPGGLRLQLRSITLTKGVLIQRNPOLCYQDTILWKDIFPHKNNQL 179

DB 123 YQK-NPSSP--DVYQVGLQLQLSLTEILSGGVYKSHNPILLCNVETINWWDIVDKTSNP 179

QY 180 ALTLIDTNRSRACHPCSPCKGRGSESSDCQSLRTVTCAGGC-ARCKGRLPTDCCHE 238

DB 180 TMLIIPHAFEQCKCHGCVNGSWAPGPGCHQCKFKLLCAEQCNRCRGPKPIDCCNE 239

QY 239 QCAAGCTGPKHSDCLACLHFNHSGICELHCPALVYNTDFTFESMPNPGRYTFGASCVTA 298

DB 240 HCAGGCTGPRATDCLACRDFNDGCTKDTCPPKIYDIVSHQVVDNPNIKYTFGAACVKE 299

QY 299 CPNYLSTDVGSCTVLCPLHNOEVTADGTQCEKSKPCAR 340

DB 300 CPSNYVTE-GACVRSACAGMLEVD-ENGKSCKEPCDGVCPK 339

RESULT 13 A27131

epidermal growth factor receptor - fruit fly (Drosophila melanogaster) (fragment)
 C:Species: Drosophila melanogaster
 C>Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 23-May-1997
 C:Accession: A27131

R:Schejter, E.D.; Segal, D.; Glazer, L.; Shilo, B.Z.
Cell 46, 1091-1101, 1986
A:Title: Alternative 5' exons and tissue-specific expression of the Drosophila EGF recep
A:Reference number: A27131; MUID:87002474; PMID:3093080
A:Accession: A27131
A:Molecule type: mRNA
A:Residues: 1-843 <SCH>
C:Genetics:
A:Gene: FlyBase:Egfr
A:Cross-references: FlyBase:PBgn0003731
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; growth factor receptor

Query Match 25.2%; Score 575.5; DB 2; Length 843;
Best Local Similarity 36.6%; Pred. No. 1.1e-31;
Matches 119; Conservative 45; Mismatches 130; Indels 31; Gaps 7;

QY 24 QVCTGDMKRLRPASPEHLDMLRHLVYOGQVQGNLELTLYLPT-NASLSFLQDIQEVQ 82
DB 51 KVICGKGLRLSPVSEHNYNLRDRYNTCTVVDNKLTLWPNENLDLSFLDIREVTG 110
QY 83 YVLIHNRVQVPLQRLRIVRGTLF-----EDNYALAVLDNGDPLNNTPVVGASPG 137
DB 111 YILISHVDVKKVFPKQLIINGRTLFSLSVSEKVALFV-----TVSKM 154
QY 138 RELQLRSITELKGGVLTORNPOLCYOQTLWKDFHKNQLALTLICTNSRACHPCSP 197
DB 155 YLLEIPDLRDVLNGVGPHNNVNLCHMETIQSEIVSNGTDAYNYDTAFERECPKCHE 214
QY 198 MCKGRSGRGESSEDCCSLTRTVACGCA--RCKGELPTDCCHSQCAAGCTGPKHSDCLAC 255
DB 215 SCTHG-CWGEKPKNQKFSKLTCSPOCAGRCYGPKEPCCHLFCAGGCTGPTOKDCIAC 273
QY 256 LFNHSGICEHLCPALVYNTDTFESMPNPEGRYTFGASCVTACPNYVLTSDVGSCTLV 315
DB 274 KNFFDEAVSKBCEPPMRKYNPTTYVLETPNPKYAYGATCVKBCP-GHLLRDNGACVRSC 332
QY 316 PLHNQEVTAEDGTQCEKSKPCAR 340
DB 333 PQDXNDKGGE-----CVPCNGPCPK 352

RESULT 14
E88257
protein let-23 (imported) - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-Aug-2001
C:Accession: E88257
R:Anonymous, The C. elegans Sequencing Consortium.
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: E88257
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1323 <STO>
A:Cross-references: GB:chr_II; PIDN:CAA93882.1; PID:G3881523; GSPDB:GN00020
C:Genetics:
A:Gene: let-23
A:Map position: 2
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

Query Match 18.3%; Score 419; DB 2; Length 1323;
Best Local Similarity 29.0%; Pred. No. 8e-21;
Matches 102; Conservative 59; Mismatches 131; Indels 60; Gaps 14;

QY 25 VCTGTDMKRLRPASPEHLDMLRHLVYOGQVQGNLELTLYLPTN----- 68
DB 39 LCSGTGNGISRYGTGNI-LEDETWRGCRVYGNLEITWENAIKWRSTNSVDPK 97
QY 69 -----ASLSFLQDIQEVQGVYVLIAPNRVQVPLQRLRIVRGTLFEDNYALAVLDNGDP 122

DB 98 NEDSPLKSNFFNLEIRGSLIYRANIOKISFPLRLVIYGVDEFDN-ALYHKNDK- 155
QY 123 LNNTPVTGASPGLELQLRSITELKGGVLTORNPOLCY-QDTILWKDIFHKNQLAL 181
DB 156 -----VHEVVRBELRVIRNGSVTIQDNFQMCYIGDKIDWKELLYDPD--VQ 199
QY 182 TLIDTNRSRACH-----PCSPWCKSGSRGWGSSSDCCSLTRTVCGAGGCARC---KGPL 231
DB 200 KVETTNSSHOCYQNGKSMAXCHESC-NDKCGSGDNDQVRVRSVCPKSCSQCFYSNSTS 259
QY 232 PTCCHBQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRYTF 291
DB 259 SYCCDSACLGCTGHPKNCIACSKVELDGICETPSRKIFNHKTGLRVFNPDCRYQN 318
QY 292 GASCVTACPNYL-STDVGSCTLV-PLHNQEVTAEDGTQCEK-SKPCAR 340
DB 319 GNHCVKCEPPELLIENDV--CVRHCSGHHYATKD--VRECEKCRSSSCPK 366

RESULT 15

S70712
protein-tyrosine kinase (EC 2.7.1.112) let-23 precursor - Caenorhabditis elegans
N:Alternate names: receptor tyrosine kinase let-23
C:Species: Caenorhabditis elegans
C:Date: 21-Apr-1997 #sequence_revision 09-May-1997 #text_change 11-Jan-2002
C:Accession: S70712; S73101; S13422; T27682
R:Sakai, T.; Koga, M.; Ohshima, Y.
J. Mol. Biol. 256, 548-555, 1996
A:Title: Genomic structure and 5' regulatory regions of the let-23 gene in the nematoc
A:Reference number: S70712; MUID:96177760; PMID:8604137
A:Accession: S70712
A:Status: nucleic acid sequence not shown

A:Molecule type: DNA
A:Residues: 1-1374 <SAK>
A:Cross-references: EMBL:D63426
A:Experimental source: strain N2
R:Koga, M.
A:Reference number: S73101
submitted to the EMBL Data Library, July 1995
A:Molecule type: DNA
A:Residues: 1-50, 'G', 52-1374 <KOG>
A:Cross-references: EMBL:D63426; NID:G1407562; PIDN:BAA09729.1; PID:G1407563
A:Experimental source: strain N2
R:Arlian, R.V.; Koga, M.; Mendel, J.E.; Ohshima, Y.; Sternberg, P.W.
Nature 348, 693-699, 1990
A:Title: The let-23 gene necessary for Caenorhabditis elegans vulval induction encodes:
A:Reference number: S13422; MUID:91080919; PMID:1979659
submitted to the EMBL Data Library, March 1996
R:Thomas, K.
A:Reference number: 220404
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 52-1374 <ARO>
A:Cross-references: EMBL:270038; PIDN:CAA93882.1; GSPDB:GN00020; CESP:ZK1067.1
A:Experimental source: clone ZK1067
C:Genetics:
A:Gene: let-23; CESP:ZK1067.1
A:Map position: 2
A:Introns: 44/1; 51/1; 90/1; 112/3; 165/3; 217/1; 290/1; 379/1; 418/1; 448/2; 565/1; 6
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-1374/Product: protein-tyrosine kinase let-23 #status predicted <MAT>
F:934-1199/Domain: protein kinase homology <KIN>
F:942-950/Region: protein kinase ATP-binding motif

Query Match

18.3%; Score 419; DB 2; Length 1374;

Best Local Similarity 29.0%; Pred. No. 8.3e-21;
Matches 102; Conservative 59; Mismatches 131; Indels 60; Gaps 14;
QY 25 VCTGDMKRLPASPETHLDMLRHLYGCGVQVQGNLELYLPTN----- 68
DB 90 LAGSITNGISRYGTGNI--LEDLETMYRGCRVYGNLEITWBIANBIKKWRESTNSTVDPK 148
QY 69 -----ASLSPLQDIQVOGVVLIHNOVQVPLQRLRIVRGTLQFEDNVALAVLNDGDP 122
DB 149 NEDSPKSLNFDNLEIRGSLIYRANIKISFPLRVYIGVEFHDN--ALYIHNDK- 206
QY 123 LNNITPVTGASPGGLRELQRLSLEILKGVLIQIQRNPOLCY-QDTILWKDIFHKNQLAL 181
DB 207 -----VHEVVMRELVRVIRNGSVTIQDNPKMCIYIGDKIDWKELLYDPP--VQ 250
QY 182 TLIDTNRSPACH-----PCSPMKGSRGNGESSEDQSLRTVTCAGGCARC---KGPL 231
DB 251 KYETTNHQHCYQNGSKMAKCHESC-NDKWSGDNDCQVRVSRVCPKSCQCFYSNSTS 309
QY 232 PFDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTPESMNPGRYTF 291
DB 310 SYECCDSACLGCTGHPKNCIACSKYELDGICIECTPSRKIFNHKTGLVNPDRQYN 369
QY 292 GASCVTACPNYL-STDVGSCTLVLC-PLHNQEVTAEDGTQRCCK-SKPCAR 340
DB 370 GNHCVKCEPELLIENDV--CVRHCSGHHYDATKD--VRECEKCRSSSCP 417
RESULT 16
S70713
protein-tyrosine kinase let-23 precursor homolog - Caenorhabditis vulgaris
N:Alternate names: receptor tyrosine kinase let-23 homolog
C:Species: Caenorhabditis vulgaris
C>Date: 21-Apr-1997 #sequence_revision 09-May-1997 #text_change 19-Dec-1997
C:Accession: S70713
R:Sakai, T.; Koga, M.; Ohshima, Y.
J. Mol. Biol. 236, 548-555, 1996
A:Title: Genomic structure and 5' regulatory regions of the let-23 gene in the nematode
A:Reference number: S70712; MUID:96177760; PMID:8604137
A:Accession: S70713
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1369 <SAK>
A:Cross-references: EMBL:D63427
C:Genetics:
A:Gene: let-23
A:Introns: 42
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein ki
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-1369/Product: protein-tyrosine kinase let-23 homolog #status predicted <MAT>
F:929-1194/Domain: protein kinase homolog <KIN>
F:937-945/Region: protein kinase ATP-binding motif
Query Match 18.2%; Score 416; DB 2; Length 1369;
Best Local Similarity 28.6%; Pred. No. 1.3e-20;
Matches 105; Conservative 59; Mismatches 135; Indels 68; Gaps 13;
QY 25 VCTGDMKRLPASPETHLDMLRHLYGCGVQVQGNLELYLPTN----- 68
DB 83 VCSGTNLLSRYSNGMI--LEDLEHMYRGCRVYGNLEITWBIANBIKKWRESTNSTQVTDAD 141
QY 69 -----ASLSPLQDIQVOGVVLIHNOVQVPLQRLRIVRGTLQFEDNVALAVLNDGDP 124
DB 142 IDYLTVPFFDLEIRGSLIYRANIKISFPLRVYIGVEFHDN--SLYHQNEK--- 197
QY 125 NITPVTGASPGGLRELQRLSLEILKGVLIQIQRNPOLCYQDT--ILWKDIFHKNQLALT 183
DB 198 -----VNLVVKELAVIRNGSVSIQNNPRMCFATKVDWNBELLYDXSRQKVE- 244
QY 184 IDTNRSPACHPCSPM-----CKSGRCWGESSEDQSLRTVTCAGGCARCCKGLPT--- 233
DB 245 -XNHHKACXWNGELIASXHENCK--DKCWGNDNDQKIYRSVCPEKSCQFYSNITQSY 302

QY 234 DCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTPESMNPGRYTFGA 293
DB 303 ECCSSCLGGCTNHGPDSCIACSKYEMDEMCICTCPARKIFNHKTGRVLPNDGRIYQGN 362
QY 294 SCVTACPNYLSLTDVGSCTLVCPHLNQEVTAEDGT---QRCEKC-SKPCARGTHSLLRP 349
DB 363 HCVKECPPELLIXND-----CVRHCSGHHYDATKDWRECEKCPSSGC-----PKI 409
QY 350 AAVPVPL 356
DB 410 CTVDGFL 416
RESULT 17
GAPPE
epidermal growth factor receptor - fruit fly (Drosophila melanogaster)
N:Contains: protein-tyrosine kinase (EC 2.7.1.12) erbB
C:Species: Drosophila melanogaster
C>Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 11-Jun-1999
C:Accession: A00640; A38021
R:Linne, E.; Glaser, L.; Segal, D.; Schlessinger, J.; Shilo, B.Z.
Cell 40, 599-607, 1985
A:Title: The Drosophila EGF receptor gene homolog: conservation of both hormone binding
A:Reference number: A00640; MUID:85124611; PMID:2982499
A:Accession: A00640
A:Molecule type: DNA
A:Residues: 1-1330 <LIV>
A:Cross-references: EMBL:X03054
R:Wadsworth, S.C.; Vincent III, W.S.; Bilodeau-Wentworth, D.
Nature 314, 178-180, 1985
A:Title: A Drosophila genomic sequence with homology to human epidermal growth factor
A:Reference number: A38021; MUID:85137938; PMID:2983232
A:Accession: A38021
A:Molecule type: DNA
A:Residues: 'A', 832-866, 'V', 868-943, 'OTPSLVK', <WAD>
A:Cross-references: EMBL:X02293; NID:97922; PIDN:CAA26157.1; PID:g929565
C:Comment: This sequence is tentative because the introns have not been identified.
C:Genetics:
A:Gene: FlyBase:Egfr
A:Cross-references: FlyBase:FBgn0003731
A:Map position: 2 57F
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosph
F:1-732/Domain: extracellular #status predicted <EXT>
F:733-764/Domain: transmembrane #status predicted <TMM>
F:765-1330/Domain: intracellular #status predicted <INT>
F:808-1072/Domain: protein kinase homology <KIN>
F:816-824/Region: protein kinase ATP-binding motif
F:122-300,324,363,518,688,695,700/Binding site: carbohydrate (Asn) (covalent) #status
F:774/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F:843/Active site: Lys #status predicted
F:1181/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi.
Query Match 18.1%; Score 415; DB 1; Length 1330;
Best Local Similarity 32.5%; Pred. No. 1.5e-20;
Matches 87; Conservative 36; Mismatches 115; Indels 30; Gaps 6;
QY 80 VQGVLIHNOVQVPLQRLRIVRGTLQF-----EDNYVALAVLNDGDPNITPTVTCASP 134
DB 38 ITNYIVGLDIPCTLSYRLQIIRGTLFSLSVSEKYLEFV-----TY 81
QY 135 GGLRELQRLSLEILKGVLIQIQRNPOLCYQDTILWKDIFHKNQLALTIDTNRSPACH 194
DB 82 SKWYTLIPLDRLVNLGVQVGHNNYLNCHMETIOKSEIVSNGTDAYNYDFTAPERCEPK 141
QY 195 CSPCKGSRGNGESSEDQSLRTVTCAGGC--RCKGPLPTDCCHEQCAAGCTGPKHSDC 252
DB 142 CHESCTHG-CWGEGRKNCQKPSKLTCSPPQAGGRCYGFKPRECCCLFCAGGCTGPTQKDC 200
QY 253 LACLFHNSGICELHCPALVYNTDTPESMNPGRYTFGASCVTACPNYLSLTDVGSCT 312
DB 201 IACKNFFDEAVSKCECPMKRYNITVLEINPESKAYGATCVKECP-GHLLRDNGACV 259

QY 313 LVCPLENQEVTAEDGTQRCCKSKPCAR 340
 |||
 Db 260 RSCPQKMDKGE-----CVPCNGCPK 282

RESULT 18

D45558
 C:Species: Schistosoma mansoni
 C:Date: 22-Apr-1993 #sequence_revision 19-May-1994 #text_change 18-Jun-1999
 C:Accession: D45558; S27839

Mol. Biochem. Parasitol. 53, 17-32, 1992

R:Shoemaker, C.B.; Ramachandran, H.; Landa, A.; dos Reis, M.G.; Stein, L.D.

A:Title: Alternative splicing of the Schistosoma mansoni gene encoding a homologue of ep

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-366/Product: epidermal growth factor receptor homolog 5 #status predicted <MAT>

A:Reference number: A45558; MUID:92365727; PMID:1501637

A:Accession: D45558

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-366 <SHO>

A:Cross-references: EMBL:M86399; NID:G160963; PIDN:AAA29869.1; PID:G160964

A:Note: sequence extracted from NCBI backbone (NCBIP:111133)

C:Genetics:

A:Gene: SR

C:Superfamily: fluke epidermal growth factor receptor homolog 1; protein kinase homology

C:Keywords: alternative splicing; ATP; receptor

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-366/Product: epidermal growth factor receptor homolog 5 #status predicted <MAT>

Query Match 15.9%; Score 363.5; DB 2; Length 366;

Best Local Similarity 29.3%; Pred. No. 1.3e-17;

Matches 106; Conservative 46; Mismatches 129; Indels 59; Gaps 13;

QY 36 PASEPETHLMRLHYOGQVQVGNLELTYP-----TNASLSFLQDIOGVGYVLIARHQ 90

|||
 Db 42 PNPKHITQTYIKFLYGGCTHIIIGLVICGLEKLENGSDPDLSEFKIEDVSGYVYVIGNS 101

|||
 QY 91 VRQVPLQRLIRIVRGTLQFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELOLSLTILK 150

|||
 Db 102 VKTISLPSLKVRG-----EPGYRMTISAALVISRNSLEILDRSLTAQR 148

|||
 QY 151 GGVLIQRNPQLC-YQDTILMKDIFHNQ-----LALTLDITNRSR- 190

|||
 Db 149 NDIVALLNNOFLCNFGFTIDWEQIFEDNRKQMEIPORKEKTVSHAGCDIALRKYTDDRTHK 208

|||
 QY 191 ACHPCSPCKG-SRCWGESSEDCQSILTR-----TVCAGGCARCKGPLPTDCCHQCAAG 243

|||
 Db 209 SCHGSCPVNGRGYCMGPKEPCQMKLKCANNPDNYCLGGR-----TTQPCLEECLOG 262

|||
 QY 244 C-TGPKHSDCLACLFHNSGICELHCPALVTYNTDFTSPMPNPEGRYTFGASCVTACPN 302

|||
 Db 263 CSTRP--GNCRACKAMNDGKVCQCPPLIVSRBSRTVANPEFKYFHDICVKNCPAP 320

|||
 QY 303 YLSTDVGSCTLYCPLHNOEVAEDGTQRCCKSK 336

|||
 Db 321 FLKSD-SYCVISCDLNTQ--IPVNGT--CKDCPK 349

RESULT 19

A45558

C:Species: Schistosoma mansoni

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A45558; S27836

Mol. Biochem. Parasitol. 53, 17-32, 1992

R:Shoemaker, C.B.; Ramachandran, H.; Landa, A.; dos Reis, M.G.; Stein, L.D.

A:Title: Alternative splicing of the Schistosoma mansoni gene encoding a homologue of ep

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-366/Product: epidermal growth factor receptor homolog 5 #status predicted <MAT>

A:Reference number: A45558; MUID:92365727; PMID:1501637

A:Accession: A45558

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1717 <SHO>

A:Cross-references: EMBL:M86396; NID:G160957; PIDN:AAA29866.1; PID:G160958

A:Note: sequence extracted from NCBI backbone (NCBIP:111129)

C:Genetics:

A:Gene: SR

C:Superfamily: fluke epidermal growth factor receptor homolog 1; protein kinase homolc

C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; phosphoprote

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-1717/Product: epidermal growth factor receptor homolog 1 #status predicted <MAT>

F:1018-1323/Domain: protein kinase homology <KIN>

F:1026-1034/Region: protein kinase ATP-binding motif

Query Match 15.9%; Score 363.5; DB 1; Length 1717;

Best Local Similarity 29.3%; Pred. No. 6.4e-17;

Matches 100; Conservative 46; Mismatches 129; Indels 59; Gaps 13;

QY 36 PASEPETHLMRLHYOGQVQVGNLELTYP-----TNASLSFLQDIOGVGYVLIARHQ 90

|||
 Db 42 PNPKHITQTYIKFLYGGCTHIIIGLVICGLEKLENGSDPDLSEFKIEDVSGYVYVIGNS 101

|||
 QY 91 VRQVPLQRLIRIVRGTLQFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELOLSLTILK 150

|||
 Db 102 VKTISLPSLKVRG-----EPGYRMTISAALVISRNSLEILDRSLTAQR 148

|||
 QY 151 GGVLIQRNPQLC-YQDTILMKDIFHNQ-----LALTLDITNRSR- 190

|||
 Db 149 NDIVALLNNOFLCNFGFTIDWEQIFEDNRKQMEIPORKEKTVSHAGCDIALRKYTDDRTHK 208

|||
 QY 191 ACHPCSPCKG-SRCWGESSEDCQSILTR-----TVCAGGCARCKGPLPTDCCHQCAAG 243

|||
 Db 209 SCHGSCPVNGRGYCMGPKEPCQMKLKCANNPDNYCLGGR-----TTQPCLEECLOG 262

|||
 QY 244 C-TGPKHSDCLACLFHNSGICELHCPALVTYNTDFTSPMPNPEGRYTFGASCVTACPN 302

|||
 Db 263 CSTRP--GNCRACKAMNDGKVCQCPPLIVSRBSRTVANPEFKYFHDICVKNCPAP 320

|||
 QY 303 YLSTDVGSCTLYCPLHNOEVAEDGTQRCCKSK 336

|||
 Db 321 FLKSD-SYCVISCDLNTQ--IPVNGT--CKDCPK 349

RESULT 20

T43220

C:Species: Schistosoma mansoni

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-May-2000

C:Accession: T43220

Mol. Endocrinol. 10, 857-866, 1996

R:Shoemaker, M.; Chan, S.J.; Steiner, D.F.

A:Title: Structure and expression of the insulin-like peptide receptor from amphioxus.

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-366/Product: epidermal growth factor receptor homolog 5 #status predicted <MAT>

A:Reference number: Z22346; MUID:96408719; PMID:8813726

A:Accession: T43220

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1363 <PAS>

A:Cross-references: EMBL:S83394; NID:G1911771; PID:G1911772; PIDN:AAB50848.1

C:Superfamily: insulin receptor; protein kinase homology

C:Keywords: hormone receptor

Query Match 15.0%; Score 342.5; DB 2; Length 1363;

Best Local Similarity 28.3%; Pred. No. 1.4e-15;

Matches 106; Conservative 40; Mismatches 124; Indels 105; Gaps 20;

QY 9 WGL-----LALLPFGAASCTQCTDMKLLPASPETHLDMRLHYOGQVQVGNLELT 63

|||
 Db 12 WAALTIVIGLGLVPSNGBEYICDSMDIRN-----VSNLRQL-ENCTVIEGYLQI- 61

|||
 QY 64 YLPTNASLSFLQDIOGVGYVLIARHQVQVP-----LQRLR-----IVRGTQ 106

|||
 Db 62 -----LLIDFAEQDYGLAPNLVEITDYPILYVRGLTNLSLFPNLAVIRCTN 112

|||
 QY 107 LPEDNYALVLDNGDPLNNTPTVTGASPGGLRELOLSLTILKGVLIQRNPOLYQDT 166


```
Db 113 LF-FNYLVWFEMLD-----NQKIGLVSLQNIITGSRVRIEKNPWLCLYDIT 156
Qy 167 ILWKDIP---HKNNQALTLIDTNRAC-HPCSPMCK-----GSRGWSESSDCQSLT 216
Db 157 IDWSFIAESGYNN---FIVDNEEBCVNFPCRCRIKHPVLQDLCAWAB--EHCQKVC 210
Qy 217 RTVCAGGACRCKGPLTDCHEQACAGCTGPKHSDCLACLFHNSGICELHCPALVTYNT 276
Db 211 PESCLGNCR-----DGISCCCHENCIGGCDGPTEDCVACKYFVNGECLIQCPDPTVOYK 266
Qy 277 D-----TFESMENPEGRY--TFGASCVTACPNYLSLTDVGS---CTLVCPHLNQEVTADG 327
Db 267 DARCITEBECENTTNSVWKLHAKCIPECPSGY-TTDINNPRLCI----- 310
Qy 328 TORCE-KCKSKPCARG 341
Db 311 --ECEGQCPKCKG 323

RESULT 21
B45558
epidermal growth factor receptor homolog precursor (splice form 2) - fluke (Schistosoma mansoni)
C:Species: Schistosoma mansoni
C>Date: 22-Apr-1993 #sequence_revision 19-May-1994 #text_change 18-Jun-1999
C:Accession: B45558; S27837
R:Shoemaker, C.B.; Ramachandran, H.; Landa, A.; dos Reis, M.G.; Stein, L.D.
Mol. Biochem. Parasitol. 53, 17-32, 1992
A:Title: Alternative splicing of the Schistosoma mansoni gene encoding a homologue of epidermal growth factor receptor homolog 4 #status predicted <SIG>
A:Reference number: A45558; MUID:92365727; PMID:1501637
A:Accession: B45558
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <SHO1>
A:Cross-references: EMBL:M86397; NID:G160959; PIDN:AAA29867.1; PID:G160960
A:Note: sequence extracted from NCBI backbone (NCBIP:111131)
C:Genetics:
A:Gene: SER
C:Superfamily: fluke epidermal growth factor receptor homolog 1; protein kinase homology
C:Keywords: alternative splicing; ATP; receptor
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-333/Product: epidermal growth factor receptor homolog 2 #status predicted <MAT>

Query Match 14.5%; Score 331; DB 2; Length 333;
Best Local Similarity 29.9%; Pred. No. 1.9e-15;
Matches 89; Conservative 40; Mismatches 115; Indels 54; Gaps 10;
Qy 36 PASPEHLDMLRLHYQCCVQVGNLELYLP-----TNASLSFLQDIQEQVGYVLIHQN 90
Db 42 PNPKHIQLTYIKFLYGGCTHIIGNLVICGLEKLENGSDPLSFLEKIEDVSGYVIQNS 101
Qy 91 VRQVPLQRLIRVGTQLFEDNALAVLDNGDPLNNTTPVTGASPGGLREQLRSLTILK 150
Db 102 VKTISPLSKVIRG-----EPGYRIMNTSAALVISRNSLEILDRLSLTAIOR 148
Qy 151 GGVLIQNPOLC-YQDTILWKDIPKRNQ-----LATLIDTNRSR- 190
Db 149 NDIVALLNQFLCNFGFTIDWEQIFEDNRKQMFIPDRKKTVSHAGCDIALRKYTDDRTKH 208
Qy 191 ACHPCSPMCKG-SRCGWSESSDCQSLTR-----TVCAGGCARCKGPLETDCHEQCAAG 243
Db 209 SCHGSCPVVNGRGYCWGPKPMCKMLKANNPDNYCLGGA-----TTQPCLEELGG 262
Qy 244 C-TGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 263 CETRP--GNCRCACKHAMNDGKVCQCPPLIVSEESRTVANPEFKYFNHFDICVKNCP 318

RESULT 22
S57245
insulin receptor (version 2) - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 02-Jun-2000
C:Accession: S57245
R:Fernandez, R.; Tabarini, D.; Azpiroz, N.; Frasch, M.; Schlessinger, J.
EMBO J. 14, 3373-3384, 1995
A:Title: The Drosophila insulin receptor homolog: a gene essential for embryonic development
A:Reference number: S57245; MUID:95354655; PMID:7628438
A:Accession: S57245
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2101 <PER>
A:Cross-references: EMBL:U28136
C:Genetics:
A:Gene: FlyBase:Inr
A:Cross-references: FlyBase:FBgn0013984
C:Superfamily: Drosophila insulin receptor; protein kinase homology
C:Keywords: ATP; receptor
F:1321-1609/Domain: protein kinase homology <KIN>
F:1329-1337/Region: protein kinase ATP-binding motif

Query Match 12.9%; Score 294; DB 2; Length 2101;
```

```
Db 113 LF-FNYLVWFEMLD-----NQKIGLVSLQNIITGSRVRIEKNPWLCLYDIT 156
Qy 167 ILWKDIP---HKNNQALTLIDTNRAC-HPCSPMCK-----GSRGWSESSDCQSLT 216
Db 157 IDWSFIAESGYNN---FIVDNEEBCVNFPCRCRIKHPVLQDLCAWAB--EHCQKVC 210
Qy 217 RTVCAGGACRCKGPLTDCHEQACAGCTGPKHSDCLACLFHNSGICELHCPALVTYNT 276
Db 211 PESCLGNCR-----DGISCCCHENCIGGCDGPTEDCVACKYFVNGECLIQCPDPTVOYK 266
Qy 277 D-----TFESMENPEGRY--TFGASCVTACPNYLSLTDVGS---CTLVCPHLNQEVTADG 327
Db 267 DARCITEBECENTTNSVWKLHAKCIPECPSGY-TTDINNPRLCI----- 310
Qy 328 TORCE-KCKSKPCARG 341
Db 311 --ECEGQCPKCKG 323

RESULT 21
B45558
epidermal growth factor receptor homolog precursor (splice form 2) - fluke (Schistosoma mansoni)
C:Species: Schistosoma mansoni
C>Date: 22-Apr-1993 #sequence_revision 19-May-1994 #text_change 18-Jun-1999
C:Accession: B45558; S27837
R:Shoemaker, C.B.; Ramachandran, H.; Landa, A.; dos Reis, M.G.; Stein, L.D.
Mol. Biochem. Parasitol. 53, 17-32, 1992
A:Title: Alternative splicing of the Schistosoma mansoni gene encoding a homologue of epidermal growth factor receptor homolog 4 #status predicted <SIG>
A:Reference number: A45558; MUID:92365727; PMID:1501637
A:Accession: B45558
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <SHO1>
A:Note: sequence extracted from NCBI backbone (NCBIP:111131)
R:Shoemaker, C.B.; Ramachandran, H.; Landa, A.; dos Reis, M.G.; Stein, L.
submitted to the EMBL Data Library, February 1992
A:Description: Alternative splicing of the Schistosoma mansoni gene encoding a homologue of epidermal growth factor receptor homolog 4 #status predicted <SIG>
A:Reference number: S27837
A:Accession: S27837
A:Molecule type: mRNA
A:Residues: 1-320, 'T', 321-333 <SHO2>
A:Cross-references: EMBL:M86397; NID:G160959; PIDN:AAA29867.1; PID:G160960
C:Genetics:
A:Gene: SER
C:Superfamily: fluke epidermal growth factor receptor homolog 1; protein kinase homology
C:Keywords: alternative splicing; ATP; receptor
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-333/Product: epidermal growth factor receptor homolog 2 #status predicted <MAT>

Query Match 14.5%; Score 331; DB 2; Length 333;
Best Local Similarity 29.9%; Pred. No. 1.9e-15;
Matches 89; Conservative 40; Mismatches 115; Indels 54; Gaps 10;
Qy 36 PASPEHLDMLRLHYQCCVQVGNLELYLP-----TNASLSFLQDIQEQVGYVLIHQN 90
Db 42 PNPKHIQLTYIKFLYGGCTHIIGNLVICGLEKLENGSDPLSFLEKIEDVSGYVIQNS 101
Qy 91 VRQVPLQRLIRVGTQLFEDNALAVLDNGDPLNNTTPVTGASPGGLREQLRSLTILK 150
Db 102 VKTISPLSKVIRG-----EPGYRIMNTSAALVISRNSLEILDRLSLTAIOR 148
Qy 151 GGVLIQNPOLC-YQDTILWKDIPKRNQ-----LATLIDTNRSR- 190
Db 149 NDIVALLNQFLCNFGFTIDWEQIFEDNRKQMFIPDRKKTVSHAGCDIALRKYTDDRTKH 208
Qy 191 ACHPCSPMCKG-SRCGWSESSDCQSLTR-----TVCAGGCARCKGPLETDCHEQCAAG 243
Db 209 SCHGSCPVVNGRGYCWGPKPMCKMLKANNPDNYCLGGA-----TTQPCLEELGG 262
Qy 244 C-TGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
Db 263 CETRP--GNCRCACKHAMNDGKVCQCPPLIVSEESRTVANPEFKYFNHFDICVKNCP 318
```


QY 207 --ESSEDQSLTRTVACGACRCKGKPLPTDCCHEQCAAGCTGPKHSD-CLACLHP--NHS 261
Db 214 KQNKAAQCCQFCNTQC--GPEGCLDGSDHICCHHECLGCSAINSTNCHACKRYIKST 271
QY 262 GICEHLHCALVYNTDTP--BSMP-----NPEGRYTFGASCVTACVYNVSLTDVGSCT 312
Db 272 QCQVSKCPR-KQYLVKFLCQSCPYWSINSTYEHYLVQGEVCTKCPVNVIS----- 323
QY 313 LVCPLHNOEVTAEDGTQCEKC 334
Db 324 -----NNQ-----TKCEKC 333

RESULT 26

A36502
Insulin receptor-related receptor precursor - guinea pig
C:Species: Cavia porcellus (guinea pig)
C:Date: 01-Feb-1991 #sequence_revision 01-Feb-1991 #text_change 23-May-1997
C:Accession: A36502
R:Shier, P.; Watt, V.M.
J. Biol. Chem. 264, 14605-14608, 1989
A:Title: Primary structure of a putative receptor for a ligand of the insulin family.
A:Reference number: A36502; MUID:89359245; PMID:2768234
A:Accession: A36502
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1300 <SH1>
A:Cross-references: GB:J05047
C:Superfamily: Insulin receptor; protein kinase homology
C:Keywords: ATP; autophosphorylation; phosphoprotein; receptor; transmembrane protein
F:977-1253/Domain: protein kinase homology <IN>
F:985-993/Region: protein kinase ATP-binding motif

Query Match 12.2%; Score 278; DB 2; Length 1300;
Best Local Similarity 27.2%; Pred. No. 3.3e-11;
Matches 106; Conservative 53; Mismatches 116; Indels 114; Gaps 26;

QY 9 WCLL--ALPPG--AASQVCTGDMKRLPASPEHLMRLHLYGQCVVQGNLSITY 64
Db 9 WCLLLVSLSSAGFNLDTNNVCPSLDIR-----SEVAELRL-ENCSVVEGHLQILL 59
QY 65 LPTNA-----SLSFLQDIOEQVYVLIHNOVRQVPLQRLR-----IVRGTLQFEDNY 112
Db 60 MFTATGCEFRSLSP-PHLTQVTDYLL-----FRVYGLESRLDLFPNLAVIRGNHFL-GY 113
QY 113 ALAVLDNGDFLNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQNPOLCQDTILMKDI 172
Db 114 ALVIFEMPH-----LRDVGLPALGAVLHGSVRVEKNQBELCHLSLTDW--- 155
QY 173 PHKNQLALTLIDTN-----RSRACHPCSPMKGS-----RCWGESSE 210
Db 156 -----GLIQTPTSTNYVGNKLGESCAVCPGTILGAAGEPCARTTFNGHTDYRCW--TSS 208
QY 211 DQSLTRTVACGACRCKGKPLPTDCCHEQCAAGCTGPKH-SDCLACLHPNHSIGICELHCP 269
Db 209 HQQRV--CPCPQGLA-C--TISGECHECLGCSQPEDPRACVACHRFYQSACHRACP 263
QY 270 ALVYNTID-----TPESMPN-----PEGRVTFG---ASCVTACPVNYSLTDVGSCTLVCP 317
Db 264 -LGVTEHESWRCVTAESCANLRSVGRASTFGIHOGLACQCPGF--TRNGS-SIFC-- 317
QY 318 HNOEVTAEDGTQCE-KCKPCARTHSL 345
Db 318 -----HKCEGLCPKCKVGTGKTI 335

RESULT 27

B47417
Insulin receptor-related receptor, secreted splice form 2 precursor - rat
N:Contains: insulin receptor-related receptor, secreted splice form 1 precursor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Feb-1994 #sequence_revision 02-Jul-1996 #text_change 20-Jun-2000
C:Accession: B47417; A47417; E41924; A41924

R:Itch, N.; Jobo, K.; Tsujimoto, K.; Ohta, M.; Kawasaki, T.
J. Biol. Chem. 268, 17983-17986, 1993
A:Title: Two truncated forms of rat insulin receptor-related receptor.
A:Reference number: A47417; MUID:93352614; PMID:7698734
A:Accession: B47417
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-469 <ITO>
A:Cross-references: GB:D13966; NID:G425790; PIDN:BAA03069.1; PID:G425791
A:Experimental source: kidney
A>Note: sequence extracted from NCBI backbone (NCBIP:136440)
A:Accession: A47417
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-361,'Y',422-469 <IT2>
A:Cross-references: GB:D13965; NID:G425788; PIDN:BAA03068.1; PID:G425789
A:Experimental source: kidney
A>Note: sequence extracted from NCBI backbone (NCBIP:136439)
R:Shier, P.; Watt, V.M.
Mol. Endocrinol. 6, 723-729, 1992
A:Title: Tissue-specific expression of the rat insulin receptor-related receptor gene.
A:Reference number: A41924; MUID:92293149; PMID:1603082
A:Contents: Sprague-Dawley
A:Accession: B41924
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 29-361,'Y',422-540 <SH1>
A:Cross-references: GB:M90661
A>Note: sequence extracted from NCBI backbone (NCBIP:106577)
A:Accession: A41924
A:Molecule type: DNA
A:Residues: 1-28 <SH2>
A:Cross-references: GB:M90660; NID:G204975; PIDN:AAA41452.1; PID:G554461
A>Note: sequence extracted from NCBI backbone (NCBIP:106574; NCBIP:106576)
C:Superfamily: insulin receptor; protein kinase homology
C:Keywords: alternative splicing; ATP; extracellular protein; receptor
F:1-540/Product: insulin receptor-related receptor, secreted splice form 2 precursor #

Query Match 11.9%; Score 272; DB 2; Length 540;
Best Local Similarity 25.4%; Pred. No. 3.3e-11;
Matches 120; Conservative 61; Mismatches 161; Indels 130; Gaps 26;

QY 1 MELAALCRWG--LLALLPPGAA--STQVCTGTDMKRLPASPEHLMRLHLYGQCVV 56
Db 1 MAVPALWPMGVYLLMSLLSGCLDTLEVCPSLDIR-----SEVTELRRL-ENCSVV 51
QY 57 QGNLE--LTYLPTN---ASLSFLQDIOEQVYVLIHNOVRQVPLQRLR-----IVRG 104
Db 52 EGHQLILMFATAAGDFRGLSFPR-LTQVTDYLL-----FRVYGLESRLDLFPNLAVIRG 106
QY 105 TOLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQNPOLCQY 164
Db 107 ARLLF-GYALLIFEMPH-----LRDIGLPSLGAVLGAVRVEKNQBELCHL 150
QY 165 DTILWKDIFHKNQLALTLIDTNRSRACHPCSPMKG-----SRCW 205
Db 151 STIDWGLL-----QPAPCANHIVGNKLGESCAVCPGLGAAGEPCVTRTFEGHTDYRCW 205
QY 206 GESSEDCQSLTRTVACGACRCKGKPLPTDCCHEQCAAGCTGPKH-SDCLACLHPNHSIGIC 264
Db 206 --TSSHQQRV--CPCPGLACTVG--GECHSCLGCSQPEDPRACVACHLYFGQVC 258
QY 265 ELHCPALVYNTDTPESMPN-----PEGRVTFG---ASCVTACPVNYSLTD----- 307
Db 259 LPACFP--GTQYESWRCVTAELCGHLREVFQGHATAFGIYEGSCLACQCPGFTRNGSSIFC 317
QY 308 ---VGSCTLVCPLENQEVTAEDGTQCEKC-----SKPCARGTHS 344
Db 318 HKCEGLCPKCKVGTGKTI--SDVQATDVLGVTHVSGSLILNRQCCVSSMRPFRPPWS 377
QY 345 LLPRPAAPVPLRMQGPAPHPVLSFLRPS-WDLVSAFYSLP-LAPLSFTSYVP 394
Db 378 HPPSLANFFLIPPLPPLP---PRLSQLPPLMT-----HLLPWAAPLLSFSSSP 420

Proc. Natl. Acad. Sci. U.S.A. 86, 114-118, 1989
A;Title: Structure of the human insulin receptor gene and characterization of its prom
A;Reference number: A32214; MUID:89098861; PMID:2911561
A;Accession: A32214
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA; mRNA
A;Residues: 1-34;1005-1382 <SE2>
A;Cross-references: GB:M23100
R;Seino, S.; Beil, G.I.
Biochem. Biophys. Res. Commun. 159, 312-316, 1989
A;Title: Alternative splicing of human insulin receptor messenger RNA.
A;Reference number: A32278; MUID:89165872; PMID:2538124
A;Accession: A32278
A;Molecule type: mRNA
A;Residues: 698-704;728-772;903-909 <SE3>
A;Cross-references: GB:M24555
A;Note: splice form 2
A;Accession: B32278
A;Molecule type: mRNA
A;Residues: 698-704;728-744;757-772;903-909 <SE4>
A;Note: splice form 1
R;Ebina, Y.; Ellis, L.; Jarnagin, K.; Ederly, M.; Graf, L.; Clauser, E.; Ou, J.; Masia;
Cell 40, 747-758, 1985
A;Reference number: A05275; MUID:85176928; PMID:2859121
A;Accession: A05275
A;Molecule type: mRNA
A;Residues: 'GGLRGVGAHTRTRGPGSRWMTAGORAPDRPRAPAA',1-170,'H',172-447,'T',44;
A;Cross-references: GB:M10051
A;Note: the authors found a long open reading frame containing two possible initiation
R;Araki, E.; Shimada, F.; Fukushima, H.; Mori, M.; Shichiri, M.; Ebina, Y.
Diabetes Res. Clin. Pract. 7(Suppl.1), S31-S33, 1989
A;Title: Characterization of the promoter region of the human insulin receptor gene.
A;Reference number: A61520; MUID:90032206; PMID:2806055
A;Contents: sequence correction
A;Accession: A61520
A;Molecule type: DNA
A;Residues: 1-33 <ARA>
A;Note: authors redetermined transcription initiation site
R;Ullrich, A.; Bell, J.R.; Chen, E.Y.; Herrera, R.; Petruzzelli, L.M.; Dull, T.J.; Gr
J.
Nature 313, 756-761, 1985
A;Reference number: A05274; MUID:85137889; PMID:2983222
A;Accession: A05274
A;Molecule type: mRNA
A;Residues: 1-744;757-899,'DT',902-1277,'N',1279-1382 <ULL>
A;Cross-references: GB:X02160
A;Note: splice form 1
R;Kostha, L.; Grako, K.; Dull, T.J.; Coussens, L.; Ullrich, A.; McClain, D.A.
EMBO J. 9, 2409-2413, 1990
A;Title: Functionally distinct insulin receptors generated by tissue-specific alternat
A;Reference number: S12553; MUID:90316094; PMID:2369896
A;Accession: S12553
A;Molecule type: mRNA
A;Residues: 744-823 <MOS>
A;Note: splice form 1, called HIR-A, has a higher affinity for insulin than splice for
A;Note: splice form 1 is expressed in adult peripheral nerve, skin, kidney, striated m
2 is predominantly expressed in adult liver
R;Fujita-Yamaguchi, Y.; Hawke, D.H.; Shively, J.E.; Choi, S.
Protein Seq. Data Anal. 1, 3-6, 1987
A;Title: Partial amino acid sequence analyses of human placental insulin receptor.
A;Reference number: S03360; MUID:88190050; PMID:3447155
A;Accession: S03360
A;Molecule type: protein
A;Residues: 194-208;347-358;582-589;595-600,'N',602-603;610-641;722-729;820-829,'E',8;
R;Lavare, J.M.; Denton, R.M.
Biochem. J. 252, 607-615, 1988
A;Title: Studies on the autophosphorylation of the insulin receptor from human placenta
A;Reference number: S02677; MUID:88326279; PMID:3166375
A;Accession: S02677
A;Molecule type: protein
A;Residues: 927-956;981-1020;1183-1195;1353-1369 <TAV>
A;Note: tyrosine autophosphorylation sites determined
R;Xu, Q.Y.; Faxton, R.J.; Fujita-Yamaguchi, Y.

RESULT 28
T43212
insulin-like growth factor 1 receptor precursor - great pond snail
N;Alternate names: molluscan insulin-related peptide receptor
C;Species: Lymnaea stagnalis (great pond snail)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C;Accession: T43212
R;Roovers, E.; Vincent, M.; van Kesteren, E.; Geraerts, W.P.M.; Planta, R.J.; Vreugdenhil
Gene 162, 181-188, 1995
A;Title: Characterization of a putative molluscan insulin-related peptide receptor.
A;Reference number: 222339; MUID:96032341; PMID:7557427
A;Accession: T43212
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1607 <ROO>
A;Cross-references: EMBL:X84994; NID:g1020139; PID:g1020140; PIDN:CAA59353.1
C;Keywords: ANP; hormone receptor; phosphotransferase; protein kinase; transmembrane p
P;1-35/Domain: signal sequence #status predicted <SIG>
F;36-1607/Product: insulin-like growth factor 1 receptor #status predicted <MAT>

Query Match 11.8%; Score 269.5; DB 2; Length 1607;
Best Local Similarity 23.1%; Pred. No. 1.6e-10;
Matches 88; Conservative 57; Mismatches 131; Indels 105; Gaps 17;

QY 25 VCTGDMKILPASPETHLMRLHYOGCVVQGNLELTLYPTNA-----SLSFLQDIQEV 80
DB 64 VCGSVDIR-----SSMONFK-LLENCTVIGSLRISLFEKALDFRHLSP-PDLREI 113
QY 81 QGYVLIHNNQVRQVPLQ-----RLRIVRGQTQFEDNYALAVLDNGDLENNTPPVIGAS 133
DB 114 TDYLLM---YRVVGLTSLKLFENLAIIRGREL-NSYALVMYEMED----- 156
QY 134 PGGURELQSLTEILKGGVLIQRNPOLCYQDTILKWDIFKNNQLALTLIDNRSBACH 193
DB 157 ---LQDGLVNLRTISRGVRLTKNFKLCVETIENWTQIGVSDPEARFT--NNKQCPN 211
QY 194 PCSPMKGSRCWG---ESSEDCQSLRTVCAGGRCARCKGPLPDCCHQCQAAGCTGPKH 249
DB 212 SKDECOQSKRWTSYDQCKGLNCKQKENTYCMEN-----GSCCHDYCLGCKVPMN 262
QY 250 SD-CLACHNHSIGELHCP-----ALVTYNTDFESMNPBEGRTYFG 292
DB 263 PDECFCKEQVFNNTCPQCPPTGYTFKFLNRCLTDXECALTNDDPNTFKLLDGEKGP 322
QY 293 ASCVTACPNYLS'DVG-----SCTLVCP--LHNQEVTAEDGTQRCCKSKPC 338
DB 323 SLCLYTPQNY---SVGSDKNKMLSOCVKRQLCPKCEHGLEINNITQDAHKLKESK-- 377
QY 339 ARGTHSLLPRAAVPVFLRMQ 359
DB 378 -----ISGPLKIQ 385

RESULT 29
INHUR
N;Contains: insulin receptor precursor splice form 1; insulin receptor precursor splice
C;Species: Homo sapiens (man)
C;Date: 05-Jun-1987 #sequence_revision 07-Jul-1995 #text_change 08-Dec-2000
C;Accession: A37348; A32214; A32278; B32278; A05275; A61520; A05274; S12553; S03360; S04
R;Seino, S.; Seino, M.; Beil, G.I.
Diabetes 39, 123-128, 1990
A;Title: Human insulin-receptor gene. Partial sequence and amplification of exons by pol
A;Reference number: A37348; MUID:91006864; PMID:2210055
A;Accession: A37348
A;Molecule type: DNA
A;Residues: 1-1382 <SEI>
A;Cross-references: GB:M32823; GB:M32824; GB:M32825; GB:M32826; GB:M32827; GB:M32828; GB
M32839; GB:M32840; GB:M32841; GB:M32842; GB:M32972; NID:G186462; PIDN:AAA59452.1; PID:G3
A;Experimental source: fetal liver
R;Seino, S.; Seino, M.; Nishi, S.; Beil, G.I.

J. Biol. Chem. 265, 18673-18681, 1990
A:Title: Substructural analysis of the insulin receptor by microsequence analyses of the presence of dithiothreitol.
A:Reference number: A36103; MUID:91009374; PMID:2211730
A:Accession: A36103
A:Molecule type: protein
A:Residues: 28-34,'X',36-44;192-195,'X',197-205,'X',299-300,'XX',303-306,'X',308-309,'PX'
R:Kasuya, J.; Paz, I.B.; Maddux, B.A.; Goldfine, I.D.; Hefta, S.A.; Fujita-Yamaguchi, Y.; Biochemistry 32, 13531-13536, 1993
A:Title: Characterization of human placental insulin-like growth factor-I/insulin hybrid A:Reference number: A54170; MUID:94079885; PMID:8257688
A:Accession: C54170
A:Molecule type: protein
A:Residues: 28-34,'X',36-38,'X',40,'X',42,'X',44-45 <KAS>
A:Experimental source: placenta
A:Note: sequence extracted from NCBI backbone (NCBIP:141174) and corrected to correspond A:Accession: D54170
A:Molecule type: protein
A:Residues: 763-765,'X',767-768,'X',770-775,'X',777-781,'X' <KA2>
A:Experimental source: placenta
A:Note: sequence extracted from NCBI backbone (NCBIP:141175) and corrected to correspond R:Schaeffer, L.; Ljungqvist, L. Biochem. Biophys. Res. Commun. 189, 650-653, 1992
A:Title: Identification of a disulfide bridge connecting the alpha-subunits of the extra A:Reference number: A44709; MUID:93112026; PMID:1472036
A:Contents: annotation; disulfide bond assignments
R:Hubbard, S.R.; Wei, L.; Ellis, L.; Hendrickson, W.A. Nature 372, 746-754, 1994
A:Title: Crystal structure of the tyrosine kinase domain of the human insulin receptor. A:Reference number: A44710; MUID:95089813; PMID:7997262
A:Contents: annotation; X-ray crystallography, 2.1 angstroms
A:Note: unphosphorylated Tyr-1189 binds Asp-1159, blocking access to the active site by site
R:Araki, E.; Shimada, F.; Uzawa, H.; Mori, M.; Ebina, Y. J. Biol. Chem. 262, 16186-16191, 1987
A:Title: Characterization of the promoter region of the human insulin receptor gene. Evidence A:Reference number: I55255; MUID:88058985; PMID:3680248
A:Accession: I55255
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-33 <RES>
A:Cross-references: GB:J03466; NID:g186466; PIDN:AAA59175.1; PID:g463119
R:McKoon, C.; Moncada, V.; Pham, T.; Salvatore, P.; Kadowaki, T.; Accili, D.; Taylor, S. Mol. Endocrinol. 4, 647-656, 1990
A:Title: Structural and functional analysis of the insulin receptor promoter.
A:Reference number: I57702; MUID:91125373; PMID:2280779
A:Accession: I57702
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-33 <RE2>
A:Cross-references: GB:M76592; NID:g186476; PIDN:AAC37604.1; PID:G553512
R:Taira, M.; Taira, M.; Hashimoto, N.; Shimada, F.; Suzuki, Y.; Kanatsuka, A.; Nakamura, Science 245, 63-66, 1989
A:Title: Human diabetes associated with a deletion of the tyrosine kinase domain of the A:Reference number: I59537; MUID:89298408; PMID:2544997
A:Accession: I59537
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1008-1123 <RE3>
A:Cross-references: GB:M27197; NID:g186466; PIDN:AAA86791.1; PID:g186468
R:Elbein, S.C. Diabetes 38, 737-743, 1989
A:Title: Molecular and clinical characterization of an insertional polymorphism of the I A:Reference number: I60112; MUID:89252471; PMID:2566545
A:Accession: I83057
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 895-1086 <RE4>
A:Cross-references: GB:M29930; NID:g186473; PIDN:AAA59177.1; PID:g186474
A:Accession: I60112
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 895-1086 <RE5>

A:Cross-references: GB:M29929; NID:g186471; PIDN:AAA59176.1; PID:g186472
C:Comment: The beta chain contains an ATP binding site, a tyrosine autophosphorylation site for myristate on lysine.
C:Comment: Autophosphorylation increases kinase activity and makes it more insulin-ind C:Genetics:
A:Gene: GDB:INSR
A:Cross-references: GDB:119352; OMIM:147670
A:Map position: 19p13.3-19p13.3
A:Introns: 34/1; 218/1; 325/1; 375/1; 423/2; 495/1; 537/1; 621/1; 677/1; 744/2; 756/2;
A:Note: abnormalities in this protein lead to type A insulin resistance syndrome, acan C:Complex: heterotetramer of 2 alpha and 2 beta chains; alpha and beta chains are deriv the two alpha chains are disulfide bonded; also naturally forms a disulfide bonded hyb C:Function:
A:Description: membrane glycoprotein that mediates the rapid metabolic and long-term g tyrosine-kinase activity; the beta chain tyrosine-kinase undergoes autophosphorylation C:Superfamily: insulin receptor; protein kinase homology
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; hormone rece eific protein kinase
F:1-1382/Product: insulin receptor precursor splice form 2 #status predicted <SF2>
F:1-744,757-1382/Product: insulin receptor precursor splice form 1 #status predicted < F:1-27/Domains: signal sequence #status predicted <SIG>
F:28-758/Product: insulin receptor alpha chain #status experimental <ALP>
F:182-339/Region: cysteine-rich
F:763-1382/Product: insulin receptor beta chain #status experimental <BET>
F:957-979/Domains: transmembrane #status predicted <TMN>
F:1021-1297/Domains: protein kinase homology <KIN>
F:1029-1037/Region: protein kinase ATP-binding motif
F:43,769,782/Binding site: carboxydrate (Asn) (covalent) #status experimental
F:52,105,138,242,282,322,364,424,445,541,633,651,698,920,933,1060/Binding site: carboh F:462-495/Disulfide bonds: #status experimental
F:551/Disulfide bonds: interchain #status experimental
F:1057,1159/Active site: Lys. Asp #status predicted
F:1185,1189,1190,1355,1361/Binding site: phosphate (Tyr) (covalent) (by autophosphoryl Query Match 11.5%; Score 263.5; DB 1; Length 1382;
Best Local Similarity 25.5%; Pred. No. 3.4e-10;
Matches 97; Conservative 51; Mismatches 130; Indels 103; Gaps 20;
QY 13 LALLPPGAAS---TQVCTGDMKLRPASPEHLDMLRLHYQGVQVQGNLELYL--- 65
DB 18 VAALLGAGHLYPGEVCPGMDIRNL-----TRL-----HELENCVTEGHQLILLMPKT 68
QY 66 -PTN-ASLSFLQDIOEVQYVLLAHNQVQVPLQRLR-----IVRQTQFEDNYALAV 116
DB 69 RPEDFRDLSPFKLIM-ITDYLL-----FRVYGLSEKDLFPNLTIVGRSLF-FNVALVI 122
QY 117 LONGDPLNTPVTCASFGGLRELQRLSLTLKGVLIQNPOLCYODTILWKDIFR-- 174
DB 123 FEVH-----LKEGLYMLMNTIGSVRIKKNELCYLATIDMSRLDSV 167
QY 175 KNNQALTLIDNRSRACHPCSPMKGS-----RCWGESSEDCQSILT 216
DB 168 EDNYVLNKDNEE-----CGDICPGTAKGKNCPTATVINGQFVERCATHS--HCQKVC 219
QY 217 RTVACGGCARKGPIPTCCHEQCAAGCTGPKH-SDCLACLFHNSGICELHCP----- 269
DB 220 PTICKSHGCTAEGL-----CCHSECLGNCSPDDPTKCVACRFYLDGRVETCPPPYVHF 275
QY 270 -----ALVTYNTDTFESMNP--EGRYTF--GASCVTACPVNYLSTD-----VGSC 311
DB 276 QDMRCVNFSCQDLHKKNRRQCHQYIVHNKKICEPCSGYTMSSNLLCTCLGPGC 335
QY 312 TLVCPHMQEVTAEDGTORCE 332
DB 336 PKVCHLLBGEKNTIDSVTSAQE 356
RESULT 30
A36080
insulin receptor precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 18-Jun-1999
C:Accession: A36080

R:Goldstein, B.J.; Dudley, A.L.
Mol. Endocrinol. 4, 235-244, 1990
A:Title: The rat insulin receptor: primary structure and conservation of tissue-specific
A:Reference number: A36080; MUID:90231337; PMID:2330003
A:Accession: A36080
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1383 <GOL>
A:Cross-references: G3:M29014; NID:G204953; PIDN:AAA41441.1; PID:G204954
A:Superfamily: insulin receptor; protein kinase homology
C:Keywords: ATP; autophosphorylation; hormone receptor; phosphoprotein; transmembrane pr
F:1022-1298/Domain: protein kinase homology <Kin>
F:1030-1038/Region: protein kinase ATP-binding motif

Query Match		11.5%; Score 263.5; DB 2; Length 1383;
Best Local Similarity		24.4%; Pred. No. 3.4e-10;
Matches 108; Conservative 58; Mismatches 151; Indels 125; Gaps 23;		
Qy	11 LLLALLPPGAAS-----TQVCTGTDMLRLPASPTHLDMLRHLYQGVVQGNLELYL 65	
Db	14 LLMAVAVAGGTAGHLYFCEVCPGMDIRNL-----TSL-----HELENCVIEGHLQILLM 64	
Qy	66 ----PTN-ASLSFLQDIQEVQVYLIAHNQVRQVPLQRLR-----IVRGTLQFEDNYA 113	
Db	65 FKTRPEDFRDLSPFKLIM-ITDYLL-----PRVYGLESKDLFENLTWIRGSRLF-FNYA 118	
Qy	114 LAVLDNGDPLNNTFTVTGASPGRLREIQLRLTEILKGVLIQRNPOLCTQDIIILWKDI- 172	
Db	119 LVIFEMVH-----LKEIGLYNLNMNITRGSVRIEKNNELCYLATIDWSRIL 163	
Qy	173 -FKNNQLALTLIDNRSRACHPCSPCKGS-----RCWGSESEDCCQ 213	
Db	164 DYEDNYIVLNKDNEE-----CGDVCFGTAKTNCPTATVINGQFVERCWTHS--HCQ 215	
Qy	214 SLRTVCAGGCARCKGPLEPTDCHEQCAAGCTGPKH-SDCLACLHFHNSGICELHCPALV 272	
Db	216 KVCPTICKSHGCTAEGL-----CCKECLGNCSEPTDPTKVCACRNFYLDGQCVCETC9PY 271	
Qy	273 TYNTD-----TF-----ESMNPGRYTFGASCVTACPNYVLTSD-----V 308	
Db	272 YHFQDWRVCNPFQODLHYKCNBRKPGCHQYVHNHKKIPECSGTYMSSNLMCTPCL 331	
Qy	309 GSCTLVCPHNRQEVTAEDGTQRCCKSKPCA-----RGTHSLLPRAAVPVPLRM 358	
Db	332 GPCPKVQCILEGKTIQSVTSQAE--LRGCTVINGSLIINIRGGNNL-----AAELEA 382	
Qy	359 QPGPAHPVLSFL--RPSWDLVS 378	
Db	383 NLGLIEEISGFLKIRSYALVS 404	

Search completed: July 4, 2004, 04:21:39
Job time : 32.4478 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2004, 04:14:33 ; Search time 20.1328 Seconds

(without alignments)
1080.456 Million cell updates/sec

Title: US-09-506-079H-12

Perfect score: 2287

Sequence: 1 MELANCRWGGLLLALLPGCA.....VGRGPDPAFVNLRYEG 419

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 90 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1878	82.1	1255	1 ERB2_HUMAN	P04626 homo sapien
2	1608.5	70.3	1257	1 ERB2_RAT	P06494 rattus norv
3	1571	68.7	1254	1 ERB2_MESAU	P06553 mesocricetu
4	796.5	34.8	703	1 EGFR_CHICK	P13387 gallus gall
5	793	34.7	1210	1 EGFR_HUMAN	P00533 homo sapien
6	789	34.5	1210	1 EGFR_MOUSE	Q01279 mus musculu
7	779.5	34.1	1308	1 ERB4_RAT	Q62956 rattus norv
8	775	33.9	1308	1 ERB4_HUMAN	Q15303 homo sapien
9	735.5	32.2	1342	1 ERB3_HUMAN	P21860 homo sapien
10	698	30.5	1339	1 ERB3_RAT	Q62799 rattus norv
11	681.5	29.8	1167	1 XMRK_XIPMA	P13388 xiphophorus
12	574.5	25.1	1426	1 EGFR_DROME	P04412 drosophila
13	419	18.3	1367	1 LT23_CAEEL	P24348 caenorhabdi
14	342.5	15.0	1363	1 ILPR_BRALA	O02466 brachiosteo
15	294	12.9	2146	1 INSR_DROME	P09208 drosophila
16	291	12.7	1477	1 HTK7_HYDAT	Q25197 hydra atten
17	290	12.7	1300	1 IRR_MOUSE	Q9WCL4 mus musculu
18	278	12.2	1300	1 IRR_CAVPO	P14617 cavia porce
19	270	11.8	581	1 IRR_RAT	Q64716 rattus norv
20	269.5	11.8	1607	1 MIRP_LYMT	Q25410 lymaea sta
21	269	11.8	1297	1 IRR_HUMAN	P14616 homo sapien
22	263.5	11.5	1382	1 INSR_HUMAN	P06213 homo sapien
23	263.5	11.5	1383	1 INSR_RAT	P15127 rattus norv
24	261	11.4	1372	1 INSR_MOUSE	P15208 mus musculu
25	258	11.3	1390	1 INSR_AEDAE	Q93105 aedes aegypt
26	257.5	11.3	1367	1 IGIR_HUMAN	P08069 homo sapien
27	252.5	11.0	1370	1 IGIR_RAT	P24062 rattus norv
28	250.5	11.0	1373	1 IGIR_MOUSE	Q60751 mus musculu
29	169	7.4	1696	1 PKCS_BRACL	Q90115 brachiosteo
30	144	6.3	1877	1 PKCS_MOUSE	Q04592 mus musculu
31	142	6.2	1959	1 AGR1_RAT	P25304 rattus norv
32	131.5	5.7	913	1 PKCS_HUMAN	Q92824 homo sapien
33	129.5	5.7	1679	1 FUR2_DROME	P30432 drosophila

ALIGNMENTS

RESULT 1	ID	ERB2_HUMAN	STANDARD;	PRT;	1255 AA.
ERB2_HUMAN	AC	P04626;			
DT	13-AUG-1987	(Rel. 05, Created)			
DT	13-AUG-1987	(Rel. 05, last sequence update)			
DT	10-OCT-2003	(Rel. 42, last annotation update)			
DE	Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)				
DE	{p185erbB2} (NEU proto-oncogene) (C-erbB-2)				
DE	surface receptor HER2 (MLN 19)				
GN	ERBB2 OR HER2 OR NGL OR NEU				

Q14162 homo sapien
P41413 rattus norv
Q63415 rattus norv
Q93038 h tumor nec
Q60846 mus musculu
Q15230 homo sapien
Q03185 giardia lam
P57999 oryctolagus
P27113 oryctolagus
Q61001 mus musculu
P27918 homo sapien
P25391 homo sapien
P29122 homo sapien
P35448 xenopus lae
Q60675 mus musculu
P98110 mus scrofa
P51559 caenorhabdi
Q61789 mus musculu
Q02858 mus musculu
P07996 homo sapien
P41990 caenorhabdi
P80012 bos taurus
Q07011 homo sapien
Q28178 bos taurus
Q03376 chironomus
Q00174 drosophila
P16109 homo sapien
P54760 homo sapien
P35441 mus musculu
Q75173 homo sapien
P28607 xenopus lae
P98095 homo sapien
Q92623 homo sapien
P19137 mus musculu
P24043 homo sapien
Q28295 canis fami
P05107 homo sapien
Q86966 homo sapien
Q28983 sus scrofa
Q9TV36 sus scrofa
P37889 mus musculu
P24014 drosophila
Q99PW7 rattus norv
P12606 xenopus lae
Q89167 bos taurus
Q89YQ4 homo sapien
Q88799 mus musculu
Q95633 homo sapien
Q9NPY3 homo sapien
Q75509 homo sapien
P18168 drosophila
Q99466 homo sapien
Q58699 methanococc
P33730 canis fami
Q61116 mus musculu
P59222 mus musculu
Q96173 homo sapien


```

FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 452 452 W -> C
FT VARIANT 654 654 /FTID=VAR_016317.
FT VARIANT 655 655 I -> V (in allele B3; dbSNP:1801201).
FT VARIANT 655 655 I -> V (in allele B2 and allele B3;
FT VARIANT 655 655 dbSNP:1801200).
FT VARIANT 1170 1170 /FTID=VAR_004078.
FT VARIANT 1170 1170 P -> A.
FT VARIANT 1170 1170 /FTID=VAR_016318.
FT SEQUENCE 1255 AA; 137909 MW; 39E9DFD404DCF962 CRC64;
Query Match 82.1%; Score 1878; DB 1; Length 1255;
Best Local Similarity 83.0%; Pred. No. 3,3e-138;
Matches 362; Conservative 9; Mismatches 45; Indels 20; Gaps 5;
QY 1 MELALACRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMLRLHYQGCQVQGNL 60
DB 1 MELALACRWGLLLALLPPGAASCTGCTDMKRLPASPETHLDMLRLHYQGCQVQGNL 60
QY 61 ELTYLPTNASLSFLQDIEQVQYVLIHNOVQVPLQLRLIRVRGTQLPEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIEQVQYVLIHNOVQVPLQLRLIRVRGTQLPEDNYALAVLDNG 120
QY 121 DPNNPTPTVGASPGGLRELQRLSUTLILKGVLIQPNPQCYQDTILKWKQIFHKNQOLA 180
DB 121 DPNNPTPTVGASPGGLRELQRLSUTLILKGVLIQPNPQCYQDTILKWKQIFHKNQOLA 180
QY 181 LFLIDNRSRACHPCSPCKSGRWGSESDQSLTRIVCAGGCAKCKGPIPTDCCHQC 240
DB 181 LFLIDNRSRACHPCSPCKSGRWGSESDQSLTRIVCAGGCAKCKGPIPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
QY 301 NYLSTDVGSCTLVCPHLNQEVASDGTQRCCKSKPCAR-----GTHSLRPPAAVPVP 355
DB 301 NYLSTDVGSCTLVCPHLNQEVASDGTQRCCKSKPCARCYGIGMEHLREVRAVTSAN 360
QY 356 LRMQPG--PAHVLSPFLRPSDLVSAFYSLPLAPLSPTSVP-----SPVSVGRGPD 405
DB 361 IQEPAGCKKIFGSLAFPLPESFDGDPASNT---APLQPEQLQVFTELEITGYLISAWPD 417
QY 406 --PDHVAVNLRYEG 419
DB 418 SLPLDSVFNQLQVIRG 433

```

```

RESULT 2
ERB2 RAT
ID ERB2 RAT STANDARD; PRT; 1257 AA.
AC P06494;
DT 01-JAN-1988 (Rel. 06, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Epidermal growth factor
DE receptor-related protein).
GN ERB2 OR NEU.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]__TaxID=10116;
RP SEQUENCE FROM N.A.
RC TISSUE=Neuroblastoma;
RX MEDLINE=86118662; PubMed=3945311;
RA Bargmann C.I., Hung M.-C., Weinberg R.A.;
RT "The neu oncogene encodes an epidermal growth factor receptor-related

```

```

protein.";
RL Nature 319:226-230 (1986).
[2]
RP SEQUENCE OF 852-905 FROM N.A.
RX TISSUE=Sciatic nerve;
RX MEDLINE=91222560; PubMed=2025425;
RA Lai C., Lemke G.;
RT "An extended family of protein-tyrosine kinase genes differentially
RT expressed in the vertebrate nervous system.";
RL Neuron 6:691-704 (1991).
[3]
RP STRUCTURE BY NMR OF 650-668.
RX MEDLINE=9215181; PubMed=1346763;
RA Gullick R.J., Bottomley A.C., Loftis F.J., Doak D.G., Mulvey D.,
RA Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;
RT "Three dimensional structure of the transmembrane region of the proto-
RT oncogenic and oncogenic forms of the neu protein.";
RL EMBO J. 11:43-48 (1992).
CC -!- FUNCTION: Essential component of a neurotrophin-receptor complex,
CC although neurotrophins do not interact with it alone. GP130 is a
CC potential ligand for this receptor. Not activated by EGF, IGF-
CC alpha and amphiregulin.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBUNIT: Heterodimer with each of the other ERBB receptors. The
CC constitutively activated oncogenic variant forms a homodimer.
CC interacts with FRK/CAP (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- PTM: Ligand-binding increases phosphorylation on tyrosine
CC residues (By similarity).
CC -!- SIMILARITY: Belongs to the EGF receptor family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: X03362; CAA27059.1; ALT_INIT.
CC PIR: A24562; TVRTNU.
CC PDB: 1IIQ; 27-JUN-01.
CC PDB: 1N8Y; 18-FEB-03.
CC InterPro: IPR000494; EGRF_L domain.
CC InterPro: IPR006211; Furin-Like.
CC InterPro: IPR006212; Furin repeat.
CC InterPro: IPR009030; Grow_fac_recep.
CC InterPro: IPR000719; Prot_kinase.
CC InterPro: IPR001245; Tyr_kinase.
CC InterPro: IPR008266; Tyr_kinase_AS.
CC InterPro: IPR004019; YLP motif.
CC Pfam: PF00757; Furin-like; 1.
CC Pfam: PF00059; Kinase; 1.
CC Pfam: PF01030; Recep_L_domain; 2.
CC Pfam: PF02757; YLP; 2.
CC PRINTS: PR00109; TYRKINASE.
CC PRODOM: PD000001; Prot_kinase; 1.
CC SMART: SM00261; FU; 4.
CC SMART: SM00219; TYRK; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
CC Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
CC Proto-oncogene; Disease mutation; 3D-structure.
CC SIGNAL 1 21 POTENTIAL.
CC CHAIN 22 1257 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
CC DOMAIN 22 654 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 655 677 POTENTIAL.
CC DOMAIN 678 1257 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 159 369 CYS-RICH.
CC DOMAIN 473 645 CYS-RICH.

```

(p185erbB2) (NEU proto-oncogene) (C-erbB-2).

ERBB2 OR NEU.

Mesocricetus auratus (Golden hamster).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;

Mesocricetus.

NCBI_TaxID=10036;

[1]

SEQUENCE FROM N.A.

TISSUE=Nerve;

MEDLINE=94193007; PubMed=7908275;

Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M.,

Yamazaki Y., Ishikawa T.;

"Cloning and activation of the Syrian hamster neu proto-oncogene.";

Gene 140:251-255(1994).

CC -!- FUNCTION: Essential component of a neu-regulin-receptor complex,

although neuregulins do not interact with it alone. GP30 is a

potential ligand for this receptor. Not activated by EGF, TGF-

alpha and amphiregulin (By similarity).

CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein

tyrosine phosphate.

CC -!- SUBUNIT: Heterodimer with each of the other ERBB receptors

(Potential). Interacts with PRKCAP (By similarity).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- PTM: Ligand-binding increases phosphorylation on tyrosine

residues.

CC -!- SIMILARITY: Belongs to the EGF receptor family.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

EMBL; D16295; BAA03801.1; -.

PIR; I48161; I48161.

HSSP; P11362; 1FGK.

InterPro; IPR000494; EGFR_L_domain.

InterPro; IPR006211; Furin-like.

InterPro; IPR005212; Furin repeat.

InterPro; IPR009030; Grow_fac_recep.

InterPro; IPR000719; Prot_kinase.

InterPro; IPR001245; Tyr_kinase.

InterPro; IPR008266; Tyr_kinase_AS.

InterPro; IPR004019; VLP_motif.

Pfam; PF00757; Furin-like; 1.

Pfam; PF00069; pkinase; 1.

Pfam; PF01030; Recep_L_domain; 2.

Pfam; PF02757; VLP_2_domain; 1.

PRINTS; PR00109; TYRKINASE.

ProDom; PD000001; Prot_kinase; 1.

SMART; SK00261; FU; 4.

SMART; SK00219; TyRK; 1.

PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;

Transferase; Glycine-protein kinase; ATP-binding; Phosphorylation;

Proto-oncogene; Disease mutation.

FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 1254 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.

FT DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 653 675 POTENTIAL.

FT DOMAIN 676 1254 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 158 368 CYS-RICH.

FT DOMAIN 472 644 CYS-RICH.

FT DOMAIN 720 987 PROTEIN KINASE.

FT NP_BIND 726 734 ATP (BY SIMILARITY).

FT BINDING 753 753 ATP (BY SIMILARITY).

FT SITE 845 845 BY SIMILARITY.

FT DISULFID 195 204 BY SIMILARITY.

DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2).

GN ERBB2 OR NEU.

OS Mesocricetus auratus (Golden hamster).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;

OC Mesocricetus.

OC NCBI_TaxID=10036;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Nerve;

RX MEDLINE=94193007; PubMed=7908275;

RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M.,

RA Yamazaki Y., Ishikawa T.;

RT "Cloning and activation of the Syrian hamster neu proto-oncogene.";

RL Gene 140:251-255(1994).

CC -!- FUNCTION: Essential component of a neu-regulin-receptor complex,

although neuregulins do not interact with it alone. GP30 is a

potential ligand for this receptor. Not activated by EGF, TGF-

alpha and amphiregulin (By similarity).

CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein

tyrosine phosphate.

CC -!- SUBUNIT: Heterodimer with each of the other ERBB receptors

(Potential). Interacts with PRKCAP (By similarity).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- PTM: Ligand-binding increases phosphorylation on tyrosine

residues.

CC -!- SIMILARITY: Belongs to the EGF receptor family.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

EMBL; D16295; BAA03801.1; -.

PIR; I48161; I48161.

HSSP; P11362; 1FGK.

InterPro; IPR000494; EGFR_L_domain.

InterPro; IPR006211; Furin-like.

InterPro; IPR005212; Furin repeat.

InterPro; IPR009030; Grow_fac_recep.

InterPro; IPR000719; Prot_kinase.

InterPro; IPR001245; Tyr_kinase.

InterPro; IPR008266; Tyr_kinase_AS.

InterPro; IPR004019; VLP_motif.

Pfam; PF00757; Furin-like; 1.

Pfam; PF00069; pkinase; 1.

Pfam; PF01030; Recep_L_domain; 2.

Pfam; PF02757; VLP_2_domain; 1.

PRINTS; PR00109; TYRKINASE.

ProDom; PD000001; Prot_kinase; 1.

SMART; SK00261; FU; 4.

SMART; SK00219; TyRK; 1.

PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;

Transferase; Glycine-protein kinase; ATP-binding; Phosphorylation;

Proto-oncogene; Disease mutation.

FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 1254 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.

FT DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 653 675 POTENTIAL.

FT DOMAIN 676 1254 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 158 368 CYS-RICH.

FT DOMAIN 472 644 CYS-RICH.

FT DOMAIN 720 987 PROTEIN KINASE.

FT NP_BIND 726 734 ATP (BY SIMILARITY).

FT BINDING 753 753 ATP (BY SIMILARITY).

FT SITE 845 845 BY SIMILARITY.

FT DISULFID 195 204 BY SIMILARITY.

DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2).

GN ERBB2 OR NEU.

OS Mesocricetus auratus (Golden hamster).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;

OC Mesocricetus.

OC NCBI_TaxID=10036;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Nerve;

RX MEDLINE=94193007; PubMed=7908275;

RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M.,

RA Yamazaki Y., Ishikawa T.;

RT "Cloning and activation of the Syrian hamster neu proto-oncogene.";

RL Gene 140:251-255(1994).

CC -!- FUNCTION: Essential component of a neu-regulin-receptor complex,

although neuregulins do not interact with it alone. GP30 is a

potential ligand for this receptor. Not activated by EGF, TGF-

alpha and amphiregulin (By similarity).

CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein

tyrosine phosphate.

CC -!- SUBUNIT: Heterodimer with each of the other ERBB receptors

(Potential). Interacts with PRKCAP (By similarity).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- PTM: Ligand-binding increases phosphorylation on tyrosine

residues.

CC -!- SIMILARITY: Belongs to the EGF receptor family.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

EMBL; D16295; BAA03801.1; -.

PIR; I48161; I48161.

HSSP; P11362; 1FGK.

InterPro; IPR000494; EGFR_L_domain.

InterPro; IPR006211; Furin-like.

InterPro; IPR005212; Furin repeat.

InterPro; IPR009030; Grow_fac_recep.

InterPro; IPR000719; Prot_kinase.

InterPro; IPR001245; Tyr_kinase.

InterPro; IPR008266; Tyr_kinase_AS.

InterPro; IPR004019; VLP_motif.

Pfam; PF00757; Furin-like; 1.

Pfam; PF00069; pkinase; 1.

Pfam; PF01030; Recep_L_domain; 2.

Pfam; PF02757; VLP_2_domain; 1.

PRINTS; PR00109; TYRKINASE.

ProDom; PD000001; Prot_kinase; 1.

SMART; SK00261; FU; 4.

SMART; SK00219; TyRK; 1.

PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

PROSITE; PS00109; PROTEIN_KINASE_TYR;

FT DISULFID 199 212 BY SIMILARITY.
 FT DISULFID 236 244 BY SIMILARITY.
 FT DISULFID 240 252 BY SIMILARITY.
 FT DISULFID 255 264 BY SIMILARITY.
 FT DISULFID 268 295 BY SIMILARITY.
 FT DISULFID 299 311 BY SIMILARITY.
 FT DISULFID 315 331 BY SIMILARITY.
 FT DISULFID 334 338 BY SIMILARITY.
 FT DISULFID 511 520 BY SIMILARITY.
 FT DISULFID 515 528 BY SIMILARITY.
 FT DISULFID 531 540 BY SIMILARITY.
 FT DISULFID 544 560 BY SIMILARITY.
 FT DISULFID 563 576 BY SIMILARITY.
 FT DISULFID 567 584 BY SIMILARITY.
 FT DISULFID 587 596 BY SIMILARITY.
 FT DISULFID 600 623 BY SIMILARITY.
 FT DISULFID 626 634 BY SIMILARITY.
 FT DISULFID 630 642 BY SIMILARITY.
 FT MOD RES 1139 1247 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD RES 1247 1247 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. -) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. -) (POTENTIAL).
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. -) (POTENTIAL).
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. -) (POTENTIAL).
 FT CARBOHYD 530 530 N-LINKED (GLCNAC. -) (POTENTIAL).
 FT CARBOHYD 571 571 N-LINKED (GLCNAC. -) (POTENTIAL).
 FT CARBOHYD 629 629 N-LINKED (GLCNAC. -) (POTENTIAL).
 FT CARBOHYD 658 658 V -> E (IN ONCOGENIC NEU).
 FT VARIANT 659 659 V -> E (IN ONCOGENIC NEU).
 SQ SEQUENCE 1254 AA; 138252 MM; 974C3791C21F2BE1 CRC64;

Query Match 68.7%; Score 1571; DB 1; Length 1254;
 Best Local Similarity 74.4%; Pred. No. 2.5e-114;
 Matches 299; Conservative 26; Mismatches 67; Indels 10; Gaps 3;

QY 1 MELALCRWGLLLALLPPGASTVCTGDMKRLPASPETHDMLRLHYQCQVQGNL 60
 DB 1 MELAAWCGGLLLALLSPGASTVCTGDMKRLPASPETHDIVRLHYQCQVQGNL 60
 QY 61 ELTYLPTNASLFTLDIOEVQGVYVIAHNVQVPLQRLIRVGRQLFEDNYALVLDNG 120
 DB 61 ELTYLPANATSLFQDIOEVQGVYVIAHNVQVPLQRLIRVGRQLFEDNYALVLDNR 120
 QY 121 DPLNNTPTWTGASGGLRLQRLSRLTEILKGVLLQVLPOLCYQDTILWKDIFPKNNOLA 180
 DB 121 DPLDNVTTATGTPGLRLQRLSRLTEILKGVLLQVLPOLCYQDTILWKDIFPKNNOLA 180
 QY 181 LTLIDNRSRACHPCSPMCKSRGCSSESQSLTRTVCAAGCARGKPLPTDCCHQC 240
 DB 181 PVDIDNRSRACHPCAPACKNHCWASPEDCQTLTGTPRAVPAARLPTDCCHQC 240
 QY 241 AGCTGPKHSDCLACHNHSGICELHCPALVYNTDTPESMPKPEGRYTCASCVTACP 300
 DB 241 ARGCTGPKHSDCLACHNHSGICELHCPALVYNTDTPESMPKPEGRYTCASCVTACP 300
 QY 301 YNYLSTDVSGCTLVCPPLHNVQVTAEDGTQRCCKSKPCAR-----GTHSLLPRAAVFVP 355
 DB 301 YNYLSTEVSGCTLVCPPLHNVQVTAEDGTQRCCKSKPCARVYGLGMEHLRGARAITSAN 360
 QY 356 LRMQQK--PAHPVLSFLRPSWDLVSAFVSLPLAPLSPTSVPI 395
 DB 361 IQEFAGCKKIFGSLAPLPSFD---GNPSSGIAPLTPESLQV 399

RESULT 4
 EGFR_CHICK STANDARD; PRT; 703 AA.
 AC P13387;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (CER)
 DE (fragment).

GN EGFR.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88261272; PubMed=3250329;
 RA Lax I., Johnson A., Howk R., Sap J., Bellot F., Winkler M.,
 RA Ullrich A., Vennstrom B., Schlessinger J., Givol D.;
 RT "Chicken epidermal growth factor (EGF) receptor: cDNA cloning,
 RT expression in mouse cells, and differential binding of EGF and
 RT transforming growth factor alpha.";
 RL Mol. Cell. Biol. 8:1970-1978(1988).
 CC -!- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,
 CC AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND
 CC VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- MISCELLANEOUS: Binding of EGF to the receptor leads to
 CC dimerization, internalization of the EGF-receptor complex,
 CC induction of the tyrosine kinase activity, stimulation of cell DNA
 CC synthesis, and cell proliferation.
 CC -!- SIMILARITY: Belongs to the EGF receptor family.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; M20386; AAA48760.1; .
 CC InterPro; IPR000494; EGFR_L domain.
 CC InterPro; IPR006211; Furin-like.
 CC InterPro; IPR006212; Furin repeat.
 CC InterPro; IPR009030; Grow_Fac_recep.
 CC InterPro; IPR00719; Prot_kinase.
 CC InterPro; IPR008266; Tyr_kinase_AS.
 CC Pfam; PF00757; Furin-like; 1.
 CC Pfam; PF01030; Recep_L_domain; 2.
 CC SMART; SM00261; FU; 4.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; PARTIAL.
 CC PROSITE; PS00109; PROTEIN_KINASE_TYR; PARTIAL.
 CC PROSITE; PS00111; PROTEIN_KINASE_DOM; PARTIAL.
 CC Transmembrane; Glycoprotein; Receptor; Signal; Transferase;
 CC Tyrosine-protein kinase; ATP-binding; Phosphorylation.
 CC SIGNAL 1 30
 CC CHAIN 31 >703 EPIDERMAL GROWTH FACTOR RECEPTOR.
 CC DOMAIN 31 654 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 655 667 POTENTIAL.
 CC DOMAIN 668 >703 CYTOPLASMIC (POTENTIAL).
 CC DISULFID 197 206 BY SIMILARITY.
 CC DISULFID 201 214 BY SIMILARITY.
 CC DISULFID 222 230 BY SIMILARITY.
 CC DISULFID 226 238 BY SIMILARITY.
 CC DISULFID 239 247 BY SIMILARITY.
 CC DISULFID 243 255 BY SIMILARITY.
 CC DISULFID 258 267 BY SIMILARITY.
 CC DISULFID 271 298 BY SIMILARITY.
 CC DISULFID 302 314 BY SIMILARITY.
 CC DISULFID 318 333 BY SIMILARITY.
 CC DISULFID 336 340 BY SIMILARITY.
 CC DISULFID 513 522 BY SIMILARITY.
 CC DISULFID 517 530 BY SIMILARITY.
 CC DISULFID 533 542 BY SIMILARITY.
 CC DISULFID 546 562 BY SIMILARITY.
 CC DISULFID 565 581 BY SIMILARITY.
 CC DISULFID 569 589 BY SIMILARITY.
 CC DISULFID 592 601 BY SIMILARITY.

FT DISULFID 605 627 BY SIMILARITY.
 FT DISULFID 630 638 BY SIMILARITY.
 FT DISULFID 634 646 BY SIMILARITY.
 FT CARBOHYD 134 134 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 190 190 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 200 200 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 359 359 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 368 368 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 420 420 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 573 573 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 578 578 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 613 613 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 633 633 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 648 648 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT NON_TER 703 703
 SQ SEQUENCE 703 AA; 77427 MW; AFP2DE11B735A690 CRC64;
 Query Match 34.8%; Score 795.5; DB 1; Length 703;
 Best Local Similarity 45.5%; Pred. No. 2.1e-54;
 Matches 155; Conservative 55; Mismatches 112; Indels 19; Gaps 7;
 QY 8 RWGLLALIPFGNA-----STQVCTGTDMLKLELPASPETHLDMLRLHYQCGVQVGNLE 61
 DB 13 RGAALVLLLLGVALCSAVEEKVCQGTNNKLTQLGHVEDHFTSLQRYNNCEVVLNLE 72
 QY 62 LTYLPTNASLFLQDIQEVQGVLLAHNOVROPVLRQLRIVRGTOLEFDNYALAVLNGO 121
 DB 73 ITVEHNRDLTLKTIQEVAGVLLALNMVDVPLENLIIRGNVLYDNSPALAVLSNYH 132
 QY 122 PLNNTPTVTCASPGGRLQLSLTEILKGGVLIQNPOLCYQDTILWKDIFKNNQAL 181
 DB 133 -MKTKY-----GLRELFPMKRLSEILNGGVKISNNPKLCNMVDVLMNDIIDTSRK-PL 182
 QY 182 TLID-TNRSRACHPSPCKGSRGSESSDQSLTRTVCAAGCA-RCKGPLPTDCCHEQ 239
 DB 183 TVLDPASNLSSCPKCHFNCTEDHCWGAGEQNCQLTKVICAQCSGRCKGVSPDCCHQ 242
 QY 240 CAGCTGPKHSCIACLHNHSGICELHCPALVTNTOTPSMENPEGRVTFGASCUTAC 299
 DB 243 CAGCTGPPRESCLACRFDRDCKTCTCPPLVLYNPTYQMDVNPBEGKYSFGATCVR 302
 QY 300 PYNLSTDVGSCTVCLPLHNQEVTAEDGTQRCCKSPCAR 340
 DB 303 PHNYVVDTRGSCVRSNTDITYEV-EENGVRKCKKCDGLCSK 342
 [1]
 RESULT 5
 EGFR_HUMAN STANDARD; PRT; 1210 AA.
 AC P00533; O00688; O00732; P06269; Q14225; Q92795; Q9B2S2; Q9GZX1;
 AC Q9H2C9; Q9H3C9; Q9H3D7; Q9UMD8; Q9UMG5;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Receptor
 DE Protein-tyrosine kinase E1B5-1).
 GN EGFR OR ERBB1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=84219729; PubMed=6328112;
 RA Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A., Tam A.W.,
 RA Lee J., Yarden Y., Liberman T.A., Schlessinger J., Downward J.,
 RA Mayes E.L.V., Whittle N., Waterfield M.D., Seeburg P.H.;
 RT "Human epidermal growth factor receptor cDNA sequence and aberrant
 RT expression of the amplified gene in A431 epidermoid carcinoma cells";
 RL Nature 309:418-425(1984).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=88217333; PubMed=3329716;
 RA Haley J.D., Whittle N., Bennett P., Kinchington D., Ullrich A.,
 RA Waterfield M.D.;
 RT "The human EGF receptor gene: structure of the 110 kb locus and
 RT identification of sequences regulating its transcription.";

RX MEDLINE=95382957; PubMed=7654368;
 RA Ilekis J.V., Stark B.C., Scoccia B.;
 RT "Possible role of variant RNA transcripts in the regulation of
 RT epidermal growth factor receptor expression in human placenta.";
 RL Mol. Reprod. Dev. 41:149-156(1995).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=97078686; PubMed=8918811;
 RA Reiter J.L., Maihle N.J.;
 RT "A 1.8 kb alternative transcript from the human epidermal growth
 RT factor receptor gene encodes a truncated form of the receptor.";
 RL Nucleic Acids Res. 24:4050-4056(1996).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=97256547; PubMed=9103388;
 RA Ilekis J.V., Gariti J., Niederberger C., Scoccia B.;
 RT "Expression of a truncated epidermal growth factor receptor-like
 RT protein (TEGFR) in ovarian cancer.";
 RL Gynecol. Oncol. 65:36-41(1997).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
 RX MEDLINE=21100872; PubMed=11161793;
 RA Reiter J.L., Threadgill D.W., Bley G.D., Strunk K.E., Danielson A.J.,
 RA Schehl Sinclair C., Peatsall R.S., Green P.J., Yee D., Lampland A.L.,
 RA Balasubramanian S., Crossley T.D., Magnuson T.R., James C.D.,
 RA Maihle N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and
 RT mouse alternative EGFR transcripts encoding truncated receptor
 RT isoforms.";
 RL Genomics 71:1-20(2001).
 RN [6]
 RP SEQUENCE OF 575-687 FROM N.A.
 RA Reiter J.L., Threadgill D.W., Danielson A.J., Schehl C.M.,
 RA Lampland A.L., Balasubramanian S., Crossley T.O., Magnuson T.R.,
 RA Maihle N.J.;
 RT "Human and mouse alternative EGFR transcripts encoding only the
 RT extracellular domain of the receptor.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 713-924 FROM N.A.
 RA Lin C.R., Chen W.S., Krueger W., Stolarsky L.S., Weber W.,
 RA Evans R.M., Verma I.M., Gill G.N., Rosenfeld M.G.;
 RT "Expression cloning of human EGF receptor complementary DNA: gene
 RT amplification and three related messenger RNA products in A431
 RT cells.";
 RL Science 224:843-848(1984).
 RN [8]
 RP SEQUENCE OF 150-962 FROM N.A.
 RX MEDLINE=84245835; PubMed=6330563;
 RA Xu Y.H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,
 RA Roe B.A., Werlino G.T., Pastan I.;
 RT "Human epidermal growth factor receptor cDNA is homologous to a
 RT variety of RNAs overproduced in A431 carcinoma cells.";
 RL Nature 309:806-810(1984).
 RN [9]
 RP SEQUENCE OF 1028-1210 FROM N.A.
 RX MEDLINE=85046483; PubMed=6093780;
 RA Simmen F.A., Gope M.L., Schulz T.Z., Wright D.A., Carpenter G.,
 RA O'Malley B.W.;
 RT "Isolation of an evolutionarily conserved epidermal growth factor
 RT receptor cDNA from human A431 carcinoma cells.";
 RL Biochem. Biophys. Res. Commun. 124:125-132(1984).
 RN [10]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE=88217333; PubMed=3329716;
 RA Haley J.D., Whittle N., Bennett P., Kinchington D., Ullrich A.,
 RA Waterfield M.D.;
 RT "The human EGF receptor gene: structure of the 110 kb locus and
 RT identification of sequences regulating its transcription.";

RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RX MEDLINE=93026370; PubMed=1408137;
RA Avivi A., Storecki K., Yaron A., Givol D.,
RT "Promoter region of the murine fibroblast growth factor receptor 2
RL (hek/KGFR) gene.";
RL Oncogene 7:1957-1962(1992).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c, and CD-1; TISSUE=Decidua, and Liver;
RX MEDLINE=93126380; PubMed=7678348;
RA Paria B.C., Das S.K., Andrews G.K., Dev S.K.;
RT "Expression of the epidermal growth factor receptor gene is regulated
RL in mouse blastocysts during delayed implantation.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:55-59(1993).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RX Hibbs M.L.;
RA Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=B6/C3; TISSUE=Liver;
RX MEDLINE=94170986; PubMed=9125255;
RA Luetteke N.C., Phillips H.K., Qiu T.H., Copeland N.G., Earp H.S.,
RX Jenkins N.A., Lee D.C.;
RT "The mouse waved-2 phenotype results from a point mutation in the EGF
RL receptor tyrosine kinase.";
RL Genes Dev. 8:399-413(1994).
[5]
RP SEQUENCE OF 1-714 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=91232866; PubMed=2030916;
RA Avivi A., Lax I., Ullrich A., Schlessinger J., Givol D., Kersse B.;
RT "Comparison of EGF receptor sequences as a guide to study the ligand
RL binding site.";
RL Oncogene 6:673-676(1991).
[6]
RP SEQUENCE OF 969-1117 FROM N.A.
RC STRAIN=C3H;
RX Eisner D.P., Serrero G.;
RA Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,
CC AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND
CC VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBUNIT: Binds RIPK1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- MISCELLANEOUS: Binding of EGF to the receptor leads to
CC dimerization, internalization of the EGF-receptor complex,
CC induction of the tyrosine kinase activity, stimulation of cell DNA
CC synthesis, and cell proliferation.
CC -!- SIMILARITY: Belongs to the EGF receptor family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X78987; CAA55587.1; -
CC EMBL; U03425; AAA17899.1; -
CC EMBL; X59698; CAA42219.1; -
CC EMBL; L06864; AAA53029.1; -
CC EMBL; Z12608; CAA78249.1; -
CC PIR; A53183; A53183.
CC HSP; P11362; 1FGK.
CC MGD; MG1:95294; Egrf.
CC GO; GO:0030339; C:endocytic vesicle; IDA.
CC GO; GO:0005622; C:intracellular; IDA.

DR GO; GO:0005515; F:protein binding; IPL.
DR InterPro: IPR000494; EGFR_L domain.
DR InterPro: IPR006211; Furin-Like.
DR InterPro: IPR006212; Furin repeat.
DR InterPro: IPR009030; Grow_fac_recep.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR008266; Tyr_kinase_AS.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00699; kinase; 1.
DR Pfam; PF01030; Recep_L domain; 2.
DR PRINTS; PD00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 5.
DR SMART; SM00219; TyRC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE TVR; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
KW Transmembrane; Glycoprotein; Receptor; Signal; Transferase;
KW Tyrosine-protein kinase; ATP-binding; Phosphorylation; Repeat.
FT SIGNAL 1 24 POTENTIAL. GROWTH FACTOR RECEPTOR.
FT CHAIN 25 1210 EPIDERMAL GROWTH FACTOR RECEPTOR.
FT DOMAIN 25 647 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 648 670 POTENTIAL.
FT DOMAIN 671 1210 POTENTIAL.
FT REPEAT 75 300 APPROXIMATE.
FT REPEAT 390 600 APPROXIMATE.
FT DOMAIN 1028 1071 SER-RICH.
FT DOMAIN 714 981 PROTEIN KINASE.
FT NP_BIND 720 728 ATP (BY SIMILARITY).
FT BINDING 747 747 ATP (BY SIMILARITY).
FT ACT_SITE 839 839 BY SIMILARITY.
FT DISULFID 190 199 BY SIMILARITY.
FT DISULFID 194 207 BY SIMILARITY.
FT DISULFID 215 223 BY SIMILARITY.
FT DISULFID 219 231 BY SIMILARITY.
FT DISULFID 232 240 BY SIMILARITY.
FT DISULFID 236 248 BY SIMILARITY.
FT DISULFID 251 260 BY SIMILARITY.
FT DISULFID 264 291 BY SIMILARITY.
FT DISULFID 295 307 BY SIMILARITY.
FT DISULFID 311 326 BY SIMILARITY.
FT DISULFID 329 333 BY SIMILARITY.
FT DISULFID 506 515 BY SIMILARITY.
FT DISULFID 510 523 BY SIMILARITY.
FT DISULFID 526 535 BY SIMILARITY.
FT DISULFID 539 555 BY SIMILARITY.
FT DISULFID 558 571 BY SIMILARITY.
FT DISULFID 562 579 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT DISULFID 595 617 BY SIMILARITY.
FT DISULFID 620 628 BY SIMILARITY.
FT DISULFID 624 636 BY SIMILARITY.
FT MOD_RES 680 680 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
FT MOD_RES 1092 1092 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1110 1110 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1172 1172 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1197 1197 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 128 128 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 413 413 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 528 528 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 568 568 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 19 19 C -> S (IN REF. 2).
FT CONFLICT 539 539 C -> W (IN REF. 5).
FT CONFLICT 991 991 L -> F (IN REF. 4).
FT CONFLICT 1116 1117 HP -> DR (IN REF. 6).
FT CONFLICT 1210 1210 AA; 134853 MW; 690E2046DF2D2F5 CRC64;
SEQUENCE

Query Match		34.5%; Score 789; DB 1; Length 1210;
Best Local Similarity		46.3%; Pred. No. 1.5e-53;
Matches 156; Conservative		43; Mismatches 116; Indels 22; Gaps 6;
QY		11 LLLALLPGAA--STQVCTGTDMKRLRLPASPEHLDMLRHLYQGVQVQGNLELTYPN 68
DB		14 LLLALCAAGGALEKKVCGTSNRLTQLGTEDHFLSLQRYNCEVVLGNLEITYVORN 73
QY		69 ASLSFLQDIQVQGVLLIAHQVQVPLQRIIVRGTLQFEDNVALAVLDNGDPLNNTTP 128
DB		74 YDLSFLKTIQVAGVLLIATNVTBRIPLNLEQIRGNALYENTVALATLSN----- 124
QY		129 VTGASPGRLQLRLSLTEILKGGVLIQIRNPOLCVQDTILWKDI-----FKHNNGLALTLI 184
DB		125 -YGNRTGLRELPRNLOEILIGAVRPSNNILCNWDTIQEDIVQVFNMSKMDL--- 180
QY		185 DTNRSRACHPSMCKGSRGWSESDQSQTRIVCAGGA-RCKGPLPTDCCHQCAAG 243
DB		181 -QSPSSCPKDPSPGNCSCWGGGEBNCKLTKIICAQCCSHRCGRSPSCCHNQCAAG 239
QY		244 CTGPKHSDCLACLFHNSGICELHCPALVYNTDTFESMPPEGRYTFGASCVTACPVNY 303
DB		240 CTGPRSDCLVQKQFQDEATCKDTCPPMLNPTTIQMDVNPBGKYSFGATCVKCKPRNY 299
QY		304 LSTDVGSCTLVCPILHNQEVTAEDGTQRCCKSPCAR 340
DB		300 VVTDHGSVCRACGPDYEV-BEDGIRKCKKCDGFCRK 335
RESULT 7		
ERBB4 RAT		STANDARD; PRT; 1308 AA.
AC		Q62956; Q922N7;
DT		15-DEC-1998 (Rel. 37, Created)
DT		28-FEB-2003 (Rel. 41, Last sequence update)
DE		28-FEB-2003 (Rel. 41, Last annotation update)
DE		Receptor protein-tyrosine kinase erBB-4 precursor (EC 2.7.1.112).
GN		ERBB4 OR TYRO-2.
OS		Rattus norvegicus (Rat).
OC		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID		10116;
[1]		
EN		SEQUENCE FROM N.A.
RP		TISSUE=Heart;
RC		MEDLINE=98221155; PubMed=9553078;
RX		Zhao Y.-Y., Sawyer D.R., Baliga R.R., Opel D.J., Han X.,
RA		Marchionni M.A., Kelly R.A.;
RT		"Neuregulins promote survival and growth of cardiac myocytes.
RT		Persistence of ErbB2 and ErbB4 expression in neonatal and adult
RT		ventricular myocytes.";
RL		J. Biol. Chem. 273:10261-10269(1998).
[2]		
RP		SEQUENCE OF 848-901 FROM N.A.
RC		TISSUE=Sciatic nerve;
RX		MEDLINE=91222560; PubMed=2025425;
RA		Lai C., Lemke G.;
RT		"An extended family of protein-tyrosine kinase genes differentially
RT		expressed in the vertebrate nervous system.";
RL		Neuron 6:691-704(1991).
[3]		
RP		SEQUENCE OF 1031-1198 FROM N.A.
RC		STRAIN=Sprague-Dawley; TISSUE=Spinal cord;
RX		MEDLINE=97184212; PubMed=9030524;
RA		Carroll S.L., Miller N.B., Froment P.W., Kim S.S., Corbett J.A.;
RT		"Expression of neuregulins and their putative receptors, ErbB2 and
RT		ErbB3, is induced during Wallerian degeneration.";
RL		J. Neurosci. 17:1642-1659(1997).
[4]		
RP		1- FUNCTION: Specifically binds and is activated by neuregulins, NRG-
RC		2. NRG-3 heparin-binding EGF-like growth factor, betacellulin and
RX		NTAK. Interaction with these factors induces cell differentiation.
RA		Not activated by EGF, TGF-A, and amphiregulin (By similarity).

CC		-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC		tyrosine phosphate.
CC		-1- SUBUNIT: Homodimer or heterodimer with each of the other ERBB
CC		receptors. Interacts with the PDZ domain of the synrophin SNTB2
CC		(By similarity).
CC		-1- SUBCELLULAR LOCATION: Type I membrane protein.
CC		-1- TISSUE SPECIFICITY: Preferentially expressed in the developing
CC		nervous system. Exhibits distinct and highly regionalized patterns
CC		of expression in the adult brain, where it is mainly found in the
CC		reticular nucleus of the thalamus. Very low levels in kidney, and
CC		heart.
CC		-1- PTM: Ligand-binding increases phosphorylation on tyrosine
CC		residues (By similarity).
CC		-1- SIMILARITY: Belongs to the EGF receptor family.
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration
CC		between the Swiss Institute of Bioinformatics and the EMBL outstation
CC		the European Bioinformatics Institute. There are no restrictions on its
CC		use by non-profit institutions as long as its content is in no way
CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC		or send an email to license@isb-sib.ch).
CC		-----
DR		EMBL; AF041838; AAC08899.1; ..
DR		EMBL; U52531; AAC53051.1; ..
DR		PIR; P0184; P0184.
DR		HSP; P11362; IFGK.
DR		InterPro; IPR000494; EGFR_L_domain.
DR		InterPro; IPR006211; Furin_Like.
DR		InterPro; IPR006212; Furin_repeat.
DR		InterPro; IPR009030; Growth_factor_recep.
DR		InterPro; IPR000719; Prot_kinase.
DR		InterPro; IPR001245; Tyr_kinase.
DR		InterPro; IPR008266; Tyr_kinase_AS.
DR		InterPro; IPR004019; YLP_motif.
DR		Pfam; PF00757; Furin-like; 1.
DR		Pfam; PF00069; Kinase; 1.
DR		Pfam; PF01030; Recept_L_domain; 2.
DR		Pfam; PF02757; FLP; 2.
DR		PRINTS; PD00109; TYRKINASE.
DR		ProDom; PD000001; Prot_kinase; 1.
DR		SMART; SM00261; FU; 5.
DR		SMART; SM00219; TyrKc; 1.
DR		PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR		PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR		PROSITE; PS00109; PROTEIN_KINASE_TIR; 1.
KW		Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
KW		Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT		SIGNAL 1 25 POTENTIAL.
FT		CHAIN 26 1308 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-4.
FT		DOMAIN 26 651 EXTRACELLULAR (POTENTIAL).
FT		TRANSMEM 652 675 POTENTIAL.
FT		DOMAIN 676 1308 CYTOPLASMIC (POTENTIAL).
FT		DOMAIN 186 334 CYS-RICH.
FT		DOMAIN 496 633 CYS-RICH.
FT		DOMAIN 718 985 PROTEIN KINASE.
FT		NP_BIND 724 732 ATP (BY SIMILARITY).
FT		BINDING 751 751 ATP (BY SIMILARITY).
FT		ACT_SITE 843 843 BY SIMILARITY.
FT		DISULFID 189 197 BY SIMILARITY.
FT		DISULFID 213 221 BY SIMILARITY.
FT		DISULFID 217 229 BY SIMILARITY.
FT		DISULFID 230 238 BY SIMILARITY.
FT		DISULFID 234 246 BY SIMILARITY.
FT		DISULFID 249 258 BY SIMILARITY.
FT		DISULFID 262 289 BY SIMILARITY.
FT		DISULFID 293 304 BY SIMILARITY.
FT		DISULFID 308 323 BY SIMILARITY.
FT		DISULFID 326 330 BY SIMILARITY.
FT		DISULFID 503 512 BY SIMILARITY.
FT		DISULFID 507 520 BY SIMILARITY.
FT		DISULFID 523 532 BY SIMILARITY.

FT DISULFID 536 552 BY SIMILARITY.
 FT FT DISULFID 555 569 BY SIMILARITY.
 FT DISULFID 559 577 BY SIMILARITY.
 FT FT DISULFID 580 589 BY SIMILARITY.
 FT DISULFID 593 614 BY SIMILARITY.
 FT FT DISULFID 617 625 BY SIMILARITY.
 FT DISULFID 621 633 BY SIMILARITY.
 FT MOD RES 1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD RES 1188 1188 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD RES 1258 1258 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD RES 1284 1284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 138 138 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 174 174 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 253 253 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 358 358 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 410 410 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 473 473 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 495 495 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 548 548 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 576 576 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 620 620 N-LINKED (GLCNAC) (POTENTIAL).
 FT CONFLICT 1062 1062 S -> N (IN REF. 3).
 FT CONFLICT 1080 1082 PPT -> SYR (IN REF. 3).
 SQ SEQUENCE 1308 AA; 146957 MW; D9445B0996A08B41 CRC64;

Query Match 34.1%; Score 779.5; DB 1; Length 1308;
 Best Local Similarity 45.9%; Pred. No. 9.2e-53;
 Matches 163; Conservative 45; Mismatches 126; Indels 21; Gaps 8;

QY 1 MEIA-ALCRWGLLL--ALLPPGAASTQVCTGTDMKLRLLPASPTHTDMLRHLYQGQGVQ 57
 DB 1 MKLATGLWWSLLVAARTVQPSASQSVCACTENKLSLSLDLQQYRALRKTYENCEVVM 60

QY 58 GNLRTVLPNASLFLDIOEQVGVVLIHNOVROVPLQRIIVRGTLFDENYALAVL 117
 DB 61 GNLRTVSIENRDLFSRSREVTGVVLVNLNPFVLPENLRIIRGTIKLYEDRYALAF 120

QY 118 DNGDPLNNTPTVTCASPGGLREIQLRLSLTEILKGGVLIQRNPOLCYQDITLWKDIFPKN 177
 DB 121 LNYRXDGNF-----GLQELGLNLTIELNGVGVVDQNKFLCYADITHWQDIVRNEM 171

QY 178 QLALTLIDNRSRACHPCSPKCGSRWGSSESDCSLRITVCAGCG-ARCKGPLPTDCC 236
 DB 172 PSNLTIVSTIGSGGCRCHKSTG-RCWGPTEHNCQLTRITVCAEQDCRCGYVYSDCC 230

QY 237 HECAAGCTGPKHSIDCLACLFHNSGICELHCPALVTYNTDTPESNPENGRYTFGASCV 296
 DB 231 HRECAAGCGSPKTDCCFACWNFNDSGACVTCQCFVFNPTTFQLEHNFNAKYTYGAPCV 290

QY 297 TACPYNVLSTDVSCSLVCLPHEVNOEVTAEQGTORCEKSKPCAR-----GTHSL 346
 DB 291 KXCPHNFV-YDSSSCVRACPSSKQEV-EENGIRMKCPCTDICKACDGTGIGTGLM 343

RESULT 8
 ERB4_HUMAN
 ID ERB4_HUMAN STANDARD; PRT; 1308 AA.
 AC Q15303;
 DT 15-DEC-1998 (Rel. 37, Created);
 DT 15-DEC-1998 (Rel. 37, Last sequence update);
 DT 10-OCT-2003 (Rel. 42, Last annotation update);
 DE Receptor protein-tyrosine kinase erbB-4 precursor (BC 2.7.1.112);
 DE (p180erbB4) (Tyrosine kinase-type cell surface receptor HER4).
 GN ERB4 OR HER4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM JM-A).
 RC TISSUE=Breast carcinoma;
 RX MEDLINE=93189574; PubMed=8383326;
 RA Plowman G.D., Culouscou J.-M., Whitney G.S., Green J.M., Carlton G.W.,

RA Poy L., Neubauer M.G., Shoyab M.;
 RT "Ligand-specific activation of HER4/p180erbB4, a fourth member of the
 RT epidermal growth factor receptor family";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:1746-1750(1993).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS JM-A AND JM-B).
 RC TISSUE=Petal brain;
 RX MEDLINE=97476287; PubMed=9334263;
 RA Elenius K., Corfas G., Paul S., Choi C.J., Rio C., Plowman G.D.,
 RA Klagsbrun M.;
 RT "A novel juxtamembrane domain isoform of HER4/erbB4. Isoform-specific
 RT tissue distribution and differential processing in response to
 RT phorbol ester";
 RL J. Biol. Chem. 272:26761-26768(1997).
 RN [3]
 RP INTERACTION WITH SNTB2.
 RX MEDLINE=20202681; PubMed=10725395;
 RA Garcia R.A., Vasudevan K., Buonanno A.;
 RT "The neuroligin receptor ErbB-4 interacts with PDZ-containing proteins
 RT at neuronal synapses";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:3596-3601(2000).
 CC -1- FUNCTION: Specifically binds and is activated by neuregulins. NRG-
 CC 2, NRG-3, heparin-binding EGF-like growth factor, betacellulin and
 CC NTAK. Interaction with these factors induces cell differentiation.
 CC Not activated by EGF, TGF-A, and amphiregulin.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: Homodimer or heterodimer with each of the other ERBB
 CC receptors (Potential). Interacts with the PDZ domain of the
 CC syntrophin SNTB2.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Comment=The 2 isoforms differ functionally in their response to
 CC phorbol ester: isoform JM-A is processed but not isoform JM-B.
 CC So, they respectively represent cleavable and noncleavable
 CC forms of the receptor. Both isoforms are expressed in
 CC cerebellum, but only the isoform JM-B is expressed in the
 CC heart;
 CC Name=JM-A;
 CC IsoId=Q15303-1; Sequence=Displayed;
 CC Name=JM-B;
 CC IsoId=Q15303-2; Sequence=VSP_002895;
 CC -1- TISSUE SPECIFICITY: Expressed at highest levels in brain, heart,
 CC kidney, in addition to skeletal muscle, parathyroid, cerebellum,
 CC pituitary, spleen, testis and breast. Lower levels in thymus,
 CC lung, salivary gland, and pancreas.
 CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine
 CC residues.
 CC -1- SIMILARITY: Belongs to the EGF receptor family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L07868; AAB59446.1; -;
 CC PIR; A47253; A47253.
 CC HSSP; P11362; 1FGK.
 CC Genew; HGNC:3432; ERBB4.
 CC MIM; 600543; -;
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.
 CC GO; GO:0004714; F:transmembrane receptor protein tyrosine kin. . . ; TAS.
 CC GO; GO:0008283; F:cell proliferation; TAS.
 CC GO; GO:0007275; P:development; TAS.
 CC InterPro; IPR00494; EGFR_L domain.
 CC InterPro; IPR006211; Furin-Like.
 CC InterPro; IPR006212; Furin repeat.
 CC InterPro; IPR009030; Grow_fac_recep.
 CC InterPro; IPR000719; Prot_kinase.

CC CC IsoId=P21860-1; Sequence=Displayed;
 CC CC Name=2; Synonyms=short form;
 CC CC IsoId=P21860-2; Sequence=VSP_002893, VSP_002894;
 CC CC -1- TISSUE SPECIFICITY: Epithelial tissues and brain.
 CC CC -1- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE
 CC CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.
 CC CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES
 CC CC AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF
 CC CC PHOSPHATIDYLINOSITOL 3-KINASE (BY SIMILARITY).
 CC CC -1- DISEASE: Overexpressed in a subset of human mammary tumors.
 CC CC -1- SIMILARITY: Belongs to the EGF receptor family.
 CC CC -----
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC CC or send an email to license@isb-sib.ch).
 CC CC -----
 CC CC ENBL; M29366; AAA35790.1; --
 CC CC ENBL; M34309; AAA35979.1; --
 CC CC ENBL; M61953; AAB26935.1; --
 CC CC PIR; A36223; A36223.
 CC CC PIR; JH0803; JH0803.
 CC CC PDB; 1M6B; 23-AUG-02.
 CC CC Genew; HGNC:3431; ERBB3.
 CC CC MIM; 190151; --
 CC CC GO; GO:0005887; C: integral to plasma membrane; TAS.
 CC CC GO; GO:0005006; F: epidermal growth factor receptor activity; TAS.
 CC CC GO; GO:0006468; P: protein amino acid phosphorylation; TAS.
 CC CC InterPro; IPR000494; EGF_R_L_domain.
 CC CC InterPro; IPR006211; Furin-like.
 CC CC InterPro; IPR006212; Furin repeat.
 CC CC InterPro; IPR009030; Grow_fac_recep.
 CC CC InterPro; IPR000719; Prot_kinase.
 CC CC InterPro; IPR001245; Tyr_kinase.
 CC CC InterPro; IPR008266; Tyr_kinase_AS.
 CC CC Pfam; PF03757; Furin-like; 1.
 CC CC Pfam; PF00069; pkinase; 1.
 CC CC Pfam; PF01030; Recep_L_domain; 2.
 CC CC PRINTS; PR00109; TYRKINASE.
 CC CC PRODOM; PD000001; Prot_kinase; 1.
 CC CC SMART; SM00261; FU; 5.
 CC CC SMART; SM00219; TyrK; 1.
 CC CC PROSITE; PS00107; PROTEIN KINASE ATP; FALSE NEG.
 CC CC PROSITE; PS00109; PROTEIN KINASE TYR; FALSE NEG.
 CC CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 CC CC Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 CC CC Alternative splicing; 3D-structure.
 CC CC SIGNAL 1 19
 CC CC FT CHAIN 20 1342
 CC CC FT DOMAIN 20 643
 CC CC FT TRANSMEM 644 664
 CC CC FT DOMAIN 665 1342
 CC CC FT DOMAIN 709 966
 CC CC FT NP_BIND 715 723
 CC CC FT BINDING 742 742
 CC CC FT ACT_SITE 834 834
 CC CC FT DISULFID 186 194
 CC CC FT DISULFID 190 202
 CC CC FT DISULFID 210 218
 CC CC FT DISULFID 214 226
 CC CC FT DISULFID 227 235
 CC CC FT DISULFID 231 243
 CC CC FT DISULFID 246 255
 CC CC FT DISULFID 259 286
 CC CC FT DISULFID 290 301
 CC CC FT DISULFID 305 320
 CC CC FT DISULFID 323 327
 CC CC FT DISULFID 500 509
 CC CC FT DISULFID 504 517
 CC CC BY SIMILARITY.

FT DISULFID 520 BY SIMILARITY.
 FT DISULFID 549 BY SIMILARITY.
 FT DISULFID 565 BY SIMILARITY.
 FT DISULFID 573 BY SIMILARITY.
 FT DISULFID 585 BY SIMILARITY.
 FT DISULFID 589 BY SIMILARITY.
 FT DISULFID 610 BY SIMILARITY.
 FT DISULFID 621 BY SIMILARITY.
 FT DISULFID 617 BY SIMILARITY.
 FT CARBOHYD 126 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 250 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 353 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 408 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 414 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 437 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 469 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 522 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 566 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 616 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARSPLIC 141
 FT
 FT VARSPLIC 184 Missing (in isoform 2).
 FT
 FT CONFLICT 560 E -> G (IN REF. 2).
 FT CONFLICT 1064 E -> G (IN REF. 2).
 FT SEQUENCE 1342 AA; 148097 MM; 7201E7F66CA374BD CRC64;
 SQ
 Query Match 32.2%; Score 735.5; DB 1; Length 1342;
 Best Local Similarity 44.0%; Pred. No. 2.5e-49;
 Matches 150; Conservative 49; Mismatches 121; Indels 21; Gaps 9;
 QY 10 GLLALLPFGAA--STQVCTGTDMLRLPASBETHLDMLRHLYOGGVQVGNLELYLPT 67
 Db 11 GLLFSLARGSEVNSQAVCGTGLNSVTDGAEHQVQTLKLYRCEVVGNGLEIVLTGH 70
 QY 68 NASLSFLQDTQEVQGVYLIARHQVQPLQRLRIVRGTLQDFEDNYALAVLDNGDPLNNTT 127
 Db 71 NADLSFLQWIREVTGVVIVAMNEFSLPLNLRVVRGTQVYDGKFAIFVM-----LNNT 125
 QY 128 PVTGASPGGLRLQLRLSLTEILLKGGVLIQBNPOLCVQDTILLWKDI FHKNNQLALTLDTN 187
 Db 126 ---NSHALRQLRLTQLTEILLGGVYIEKNDKLCMDTIDMRDVRDRD---AAILVKD 178
 QY 188 RSRACHPCSPMCKSGSCWGESSEDCQSLTRTVCCAGGC-ARCKGLPTDCCHEQCAAGCTG 246
 Db 179 NGRSCPPEVCKG-RWGPGESEDCQTLTKTICAPQCNHGHCFCGFENPQCCHDCAGGCSG 237
 QY 247 PKHSDCLACLFNHSIGICELHCPALVTYNTDTPESMNPBGRYTFGASCVTACPKYVLST 306
 Db 238 PQDTDCFACHFNDSGACVPRCPQPLVYKLTLPQLEPNPHTKYQYGGVCVASCPHNFV-V 296
 QY 307 DVGSCITLVCPLHNOEYTAEDGTQRCCKSKPCAR---GTHS 344
 Db 297 DQTSVVRACPPDKQNEVD-KNGLKQCFPGCLCPKACEGTGS 336
 RESULT 10
 ERBB3 RAT
 ID ERBB3_RAT STANDARD; PRT; 1339 AA.
 AC Q62799; Q62955;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Receptor protein-tyrosine kinase erbb3-3 precursor (EC 2.7.1.112)
 DE (c-erbB3).
 GN ERBB3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.

RC MEDLINE=96096535; PubMed=8522190;
RA Hellyer N.J., Kim H.-H., Greaves C.H., Sierke S.L., Koland J.G.;
RT "Cloning of the rat ErbB3 cDNA and characterization of the
RL recombinant protein.";
RL Gene 165:279-284(1995).
RW [2]
RP REVISIONS TO 85: 513 AND 565.
RA Hellyer N.J., Koland J.G.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RW [3]
RN SEQUENCE OF 922-1097 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Sciatic nerve;
RX MEDLINE=97184212; PubMed=9030624;
RA Carroll S.L., Miller M.L., Grohert P.W., Kim S.S., Corbett J.A.;
RT "Expression of neurogulin and their putative receptors, ErbB2 and
RT ErbB3, is induced during Wallerian degeneration.";
RL J. Neurosci. 17:1642-1659(1997).
CC -!- FUNCTION: BINDS AND IS ACTIVATED BY NEUREGULINS AND NTAK.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
CC (POTENTIAL).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE
CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.
CC -!- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES
CC AND PROMOTES ITS ASSOCIATION WITH THE PBS SUBUNIT OF
CC PHOSPHATIDYLINOSITOL 3-KINASE.
CC -!- SIMILARITY: Belongs to the EGF receptor family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U29339; AAC28498.2; .
DR EMBL; U52530; AAC53050.1; .
DR HSPB; P11362; LFCK.
DR InterPro; IPR000494; EGF_L domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow_fac_recep.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF01030; Recep_L domain; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 5.
DR SMART; SM00219; TyRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
KW transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT SIGNAL 1 19
FT CHAIN 20 1339
FT DOMAIN 20 543
FT TRANSMEM 644 562
FT DOMAIN 663 1339
FT DOMAIN 183 259
FT DOMAIN 707 964
FT NP BIND 713 721
FT BINDING 740 740
FT ACT SITE 832 832
FT ACT SITE 186 194
FT DISULFID 196 202
FT BY SIMILARITY.

FT DISULFID 210 218 BY SIMILARITY.
FT DISULFID 214 226 BY SIMILARITY.
FT DISULFID 227 235 BY SIMILARITY.
FT DISULFID 231 243 BY SIMILARITY.
FT DISULFID 246 255 BY SIMILARITY.
FT DISULFID 259 286 BY SIMILARITY.
FT DISULFID 290 301 BY SIMILARITY.
FT DISULFID 305 320 BY SIMILARITY.
FT DISULFID 323 327 BY SIMILARITY.
FT DISULFID 500 509 BY SIMILARITY.
FT DISULFID 504 517 BY SIMILARITY.
FT DISULFID 520 529 BY SIMILARITY.
FT DISULFID 533 549 BY SIMILARITY.
FT DISULFID 556 573 BY SIMILARITY.
FT DISULFID 576 585 BY SIMILARITY.
FT DISULFID 589 610 BY SIMILARITY.
FT DISULFID 613 621 BY SIMILARITY.
FT DISULFID 617 629 BY SIMILARITY.
FT CARBOHYD 126 126 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 250 250 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 414 414 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 469 469 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 522 522 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 566 566 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 1028 1028 L -> P (IN REF. 3).
SQ SEQUENCE 1339 AA; 147545 MW; 0AA5F2402BBD1B CRC64;

Query Match 30.5%; Score 698; DB 1; Length 1339;
Best local similarity 42.1%; Pred. No. 2.1e-46;
Matches 147; Conservative 46; Mismatches 130; Indels 26; Gaps 10;

Qy 3 LAALCRGLLLALLPPGAA--STVCTGTDMKLRLPASPETHLDMLRLHLVQGVQGN 59
Db 7 LQVLC-----FLSLARSEMGNSQAVCPGTGLNGSLVGTGDADNQVQTLKLYKECVYMG 62

Qy 60 LEITLPTNASLFLQDIQVQGVYLAHNOVQVPLQRLIRVGTQTLFEDNYALAVLDN 119
Db 63 LEIVLTGHADLSFLQWIREVTGTVLVANNEFSLPLNLRVVRGTQVYDGKFAIFVM-- 120

Qy 120 GDPANNTTPTVGASPGGLREILQRLSLTEILKGGVLIQNPOLCVQDTILKWDIFHKNNQL 179
Db 121 ---LNYNT---NSSHALRQLKFTLTLSLGGVYIENKDKLCHMDIHDRIIVVR--- 170

Qy 180 ALTLIDTNRGRACHPCSPCKSGRSGESSEDCQSLTRITVCAGGC-ARCKGPLPTDCHE 238
Db 171 GAETVVKNGANGANGPPCHEVCKG-RCWGPDPDQILTKITCAPQCNCRGCFGNPNQCC 229

Qy 239 QCRAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNEPEGYTGACSVTA 298
Db 230 ECAGGCGSPQDTDFACRRFNDGACVPRCPPELVVYNKLTFFLEPNHTIYQGVGVAS 289

Qy 299 CPVNYLSTVGSCTVLCPLHNQEVTAEDGTQRCCKSKPCAR---GTHS 344
Db 290 CPNHFV-VDQTFVCRACPPDKMEVD-KHGLKMCPECGGLCPKACEGTGS 336

RESULT 11
XMRK XIPMA
ID XMRK XIPMA STANDARD; PRT; 1167 AA.
AC P13388;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Melanoma receptor protein-tyrosine kinase precursor (EC 2.7.1.112).
GN XMRK OR TU.
OS Xiphophorus maculatus (Southern platyfish).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;

Cyprinodontiformes; Poeciliidae; Xiphophorus.
NCBI_TaxID=8083;
(1)
SEQUENCE FROM N.A.
MEDLINE=90051546; Pubmed=2797166;
Wittbrodt J., Adam D., Malitschke B., Mauele W., Raulf F.,
Telling A., Robertson S.M., Scharf M.;
"Novel putative receptor tyrosine kinase encoded by the melanoma-
inducing Tu locus in Xiphophorus";
Nature 341:415-421 (1989).
(2)
REVISION TO 515.
Scharf M.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROBABLE RECEPTOR WITH PROTEIN-TYROSINE KINASE ACTIVITY.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DISEASE: Involved in pigment cells malignant melanomas.
CC -!- SIMILARITY: Belongs to the EGF receptor family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X16891; CAA34770.2; -
CC PIR; S06142; S06142.
CC HSP; P11362; IFKG.
DR InterPro; IPR000494; EGF_R_L domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Growth fac recep.
DR InterPro; IPR000713; Prot kinase.
DR InterPro; IPR001245; Tyr kinase.
DR InterPro; IPR008266; Tyr kinase_AS.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; Kinase; 1.
DR Pfam; PF01030; Recep_L domain; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 5.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Transmembrane; Glycoprotein; Receptor; Signal; Transferase;
KW Tyrosine-protein kinase; ATP-binding; Phosphorylation; Proto-oncogene.
FT SIGNAL 1 25
FT CHAIN 26 1167 MELANOMA RECEPTOR PROTEIN-TYROSINE
FT KINASE.
FT DOMAIN 26 642
FT TRANSMEM 643 665
FT DOMAIN 666 1167
FT DOMAIN 710 977
FT NP_BIND 716 724
FT BINDING 743 743
FT ACT_SITE 835 835
FT DISULFID 195 204
FT DISULFID 199 212
FT DISULFID 220 228
FT DISULFID 224 236
FT DISULFID 237 245
FT DISULFID 241 253
FT DISULFID 256 265
FT DISULFID 269 296
FT DISULFID 300 311
FT DISULFID 315 330
FT DISULFID 333 337
FT DISULFID 504 513

FT DISULFID 508 521 BY SIMILARITY.
FT DISULFID 524 533 BY SIMILARITY.
FT DISULFID 537 553 BY SIMILARITY.
FT DISULFID 556 569 BY SIMILARITY.
FT DISULFID 560 577 BY SIMILARITY.
FT DISULFID 593 615 BY SIMILARITY.
FT DISULFID 618 626 BY SIMILARITY.
FT DISULFID 622 634 BY SIMILARITY.
FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 201 201 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 417 417 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 576 576 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 129934 MW; 479380749DC1D55A CRC64;
Query Match 29.8%; Score 681.5; DB 1; Length 1167;
Best Local Similarity 42.4%; Pred. No. 3.4e-45;
Matches 145; Conservative 46; Mismatches 136; Indels 15; Gaps 8;
QY 4 AALCEWGLLALLPPCAST---OVCTGDMKRLPASPETHLDMLRLHYQCCVQGN 59
DB 8 AALQ---LLLVLISRCSTDPKVCQSTNQMTK---LDNHYLKKMKWYSGCNVLEN 62
QY 60 LELTYLPTNASLFIQDIQEVQYVLIHNRQVPLQRLRIVRGTLFEDNYALAVLDN 119
DB 63 LEITYTQENQLSFLQSIQEVGYVLIANNEVSTIPLVNLRLIRGNLYEGNFTLLVMSN 122
QY 120 GDPLANNTPVTGASPGGLRELQLRSITELKGVLIQENPOLCYODTLWKDIEHKNQL 179
DB 123 YQK-NPSSP--DVYQVGLKQLSLNITELSGVKVSHNPLLCNVETINWWDIVDKTNP 179
QY 180 ALTLIDTNRSRACHPCSPMKSCRCWGESSEDCQSLTRTVACGGC-ARCKGLPTDCCHE 238
DB 180 TMNLIHAFERQCKQCDHGVNCSWAPQPGHCQKFTKLLCAEQCNRRCKGPKIDCCNE 239
QY 239 QCAAGCTGPKHSDCLACHFNHSGICELCPALVTYNTDTFESMNPBGRVTFGASCTA 298
DB 240 HCAGGCTGPRATVDCACRDNDGDKCTCPPPKIYDIVSHQVNDPNPKITFGAACVKE 299
QY 299 CPYNYLSTDVSGCTLVCPHLNQEVTAEDGTQRCCKSRPCAR 340
DB 300 CFSNVVTE-GACVRSKCSAGMLEVD-ENGKRSCKPCDGVCPK 339
RESULT 12
EGFR_DROME
ID EGFR_DROME STANDARD; PRT; 1426 AA.
AC P04412; O18370; O61601; P81868; Q9W2G0;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Epidermal growth factor receptor precursor (EC 2.7.1.112) [Bgfr]
DE (Gurken receptor) (Torpedo protein) (Drosophila relative of ERBB).
GN EGFR OR TOP C-ERBB OR DER OR CG10079.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS TYPES I AND II).
RX MEDLINE=94350209; Pubmed=8070664;
RA Clifford R., Schupbach T.;
RT "Molecular analysis of the Drosophila EGF receptor homolog reveals
RT that several genetically defined classes of alleles cluster in
RT subdomains of the receptor protein.";
RL Genetics 137:531-550(1994).
RN [2]

RP REVISIONS.
RA Clifford R., Schubach T.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=85124611; PubMed=2982499;
RA Livneh E., Glazer L., Segal D., Schlesinger J., Shilo B.-Z.;
RT "The Drosophila EGF receptor gene homolog: conservation of both
RT hormone binding and kinase domains.";
RL Cell 40:599-607(1985).
RN [4]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.
RX STRAIN=Oregon-R; Tissue=Embryo;
RX MEDLINE=87002474; PubMed=303080;
RA Schejter E.D., Segal D., Glazer L., Shilo B.-Z.;
RT "Alternative 5' exons and tissue-specific expression of the
RT Drosophila EGF receptor homolog transcripts.";
RL Cell 46:1091-1101(1986).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS TYPE I AND TYPE II), TISSUE SPECIFICITY,
RP AND MUTATION ANALYSIS.
RX MEDLINE=99102120; PubMed=9882502;
RA Lesokhin A.M., Yu S.-Y., Katz J., Baker N.E.;
RT "Several levels of EGF receptor signaling during photoreceptor
RT specification in wild-type, Ellipse, and null mutant Drosophila.";
RL Dev. Biol. 205:129-144(1999).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM TYPE I).
RX STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan X.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agayani A., An H.-J., Andrews-Franckoch C., Baldwin D.,
RA Balow R.M., Basu A., Bakendell J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchar M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flocker C., Gabrielian A.E., Garg N.S., Galbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Lasko P., Matti B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Rehnert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sider-Klanos I., Stapsom M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [7]
RP SEQUENCE OF 959-1078 FROM N.A.
RX STRAIN=Daekwanryong;
RX MEDLINE=85137936; PubMed=2983232;
RA Wadsworth S.C., Vincent W.S. III, Bilodeau-Wentworth D.;

RT "A Drosophila genomic sequence with homology to human epidermal
RT growth factor receptor.";
RL Nature 314:178-180(1985).
RN [8]
RP SEQUENCE OF 1069-1121 FROM N.A.
RX TISSUE=Embryo;
RX MEDLINE=98401146; PubMed=97311193;
RA Oates A.C., Wellberg P., Achen M.G., Wilks A.F.;
RT "Sampling the genomic pool of protein tyrosine kinase genes using the
RT polymerase chain reaction with genomic DNA.";
RL Biochem. Biophys. Res. Commun. 249:660-667(1998).
RN [9]
RP SEQUENCE OF 1133-1137; 1155-1167 AND 1209-1216 FROM N.A., AND MUTATION
RP ANALYSIS.
RX MEDLINE=92038942; PubMed=1936959;
RA Raz E., Schejter E.D., Shilo B.-Z.;
RT "Interallelic complementation among DER/flb alleles: implications for
RT the mechanism of signal transduction by receptor-tyrosine kinases.";
RL Genetics 129:191-201(1991).
RN [10]
RP REVIEW.
RX MEDLINE=97248481; PubMed=9094709;
RA Perrimon N., Perkins L.A.;
RT "There must be 50 ways to rule the signal: the case of the Drosophila
RT EGF receptor.";
RL Cell 89:13-16(1997).
CC -!- FUNCTION: BINDS TO FOUR LIGANDS: SPITZ, GURKEN, VEIN AND ARGOS,
CC WHICH IS AN ANTAGONIST. TRANSDUCES THE SIGNAL THROUGH THE RAS-RAF-
CC MAPK PATHWAY. INVOLVED IN A MYRIAD OF DEVELOPMENTAL DECISIONS.
CC CRITICAL FOR THE PROLIFERATION OF IMAGINAL TISSUES, AND FOR THE
CC DETERMINATION OF BOTH THE ANTERO-POSTERIOR AND DORSO-VENTRAL
CC POLARITIES OF THE OOCYTE. IN THE EMBRYO, PLAYS A ROLE IN THE
CC ESTABLISHMENT OF VENTRAL CELL FATES, MAINTENANCE OF AMNIOSEROSA
CC AND VENTRAL NEUROECTODERMAL CELLS, GERM BAND RETRACTION, CELL FATE
CC SPECIFICATION IN THE CENTRAL NERVOUS SYSTEM AND PRODUCTION OF
CC CUTICLE.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Type I;
CC IsoId=904412-1; Sequence=Displayed;
CC Name=Type II;
CC IsoId=904412-2; Sequence=VSP_002937;
CC -!- TISSUE SPECIFICITY: UNEXPECTEDLY EXPRESSED IN EMBRYOS. IN LARVAE,
CC UNIFORM EXPRESSION IS SEEN IN WING DISKS, GENITAL DISK, ANLAGEN OF
CC TESTIS AND OVARY, AND BRAIN CORTEX. IN EYE-ANTENNA DISK, HIGHEST
CC EXPRESSION IS ANTERIOR TO MORPHOGENETIC FURROW. LEVELS REMAIN HIGH
CC IN PHOTORECEPTOR PRECURSOR CELLS. THIS PATTERN IS REVERSED IN
CC POSTERIOR EYE DISK. IN ADULTS EXPRESSION IS HIGH IN BRAIN CORTEX
CC AND THORACIC AND ABDOMINAL GANGLIA.
CC -!- SIMILARITY: Belongs to the EGF receptor family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. See <http://www.ebi.ac.uk/ebis-sib.ch/announcement/>
CC or send an email to license@ebi.ac.uk.
CC -----
CC EMBL; AF052754; AAC08536.1; -;
CC EMBL; AF052753; AAC08536.1; JOINED.
CC EMBL; AF052754; AAC08535.1; -;
CC EMBL; AF052752; AAC08535.1; JOINED.
CC EMBL; K03054; AAA51462.1; -;
CC EMBL; K03417; AAA51460.1; -;
CC EMBL; K03416; AAA50965.1; -;
CC EMBL; AF109077; AAD26134.1; -;
CC EMBL; AF109078; AAD26132.1; -;
CC EMBL; AF109082; AAD26132.1; JOINED.
CC EMBL; AF109078; AAD26133.1; -;

DR EMBL; AF109084; AAD261133.1; JOINED.
DR EMBL; AF109079; AAD261130.1; -
DR EMBL; AF109081; AAD261130.1; JOINED.
DR EMBL; AF109079; AAD261131.1; -
DR EMBL; AF109083; AAD261131.1; JOINED.
DR EMBL; AF109080; AAD261135.1; -
DR EMBL; AE003454; AAF46732.1; -
DR EMBL; X02293; CAA26157.1; -
DR EMBL; AJ002912; CAA05747.1; -
DR EMBL; X78920; CAA55523.1; -
DR EMBL; X78918; CAA55521.1; -
DR EMBL; X78919; CAA55522.1; -
DR PIR; AC0640; GQFFE.
DR HSP; P11362; IFGK.
DR FlyBase; FBgn0003731; Egfr.
DR GO; GO:0007469; P:antennal morphogenesis; NAS.
DR GO; GO:0006916; P:anti-apoptosis; NAS.
DR GO; GO:0030381; P:eggshell pattern formation (sensu Insecta); IGI.
DR GO; GO:0007456; P:eye morphogenesis (sensu Drosophila); IMP.
DR GO; GO:0007390; P:germ-band shortening; IMP.
DR GO; GO:0007444; P:imaginal disc development; IMP.
DR GO; GO:0007479; P:leg disc proximal/distal pattern formation; IMP.
DR GO; GO:0008071; P:maternal determination of dorsal/ventral axis; IMP.
DR GO; GO:0007477; P:notum morphogenesis; IMP.
DR GO; GO:0007314; P:ovocyte anterior/posterior axis determination; NAS.
DR GO; GO:0045468; P:regulation of R8 spacing; NAS.
DR GO; GO:0016330; P:second mitotic wave (sensu Drosophila); IMP.
DR GO; GO:0007476; P:wing morphogenesis; IMP.
DR InterPro; IPR000494; EGFR_L domain.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR003030; Grow_Fac_recep.

Query Match 25.1%; Score 574.5; DB 1; Length 1426;
Best Local Similarity 36.3%; Pred. No. 8.9e-37;
Matches 118; Conservative 46; Mismatches 130; Indels 31; Gaps 7;

QY 24 QVCTGTDMKRLPASPETHLDMLRLHLYQGVQVQGNLELYLPT-NASLSFLQDIOBVOG 82
DB 100 KICIGTSRLSPVSKHHEHYNLRDRYNTCTYVDGNLKLTLPLNENLDLSFLDNIREYTG 159
QY 83 YVLIHNOVQVPLQRIURVGTOLP-----EDNYALAVLDNGDPLNNTTPTVTGASPGGL 137
DB 160 YLISHVDVKKVWPVKLQIRGLRFLSLVSEERKALFV-----TVSKM 203
QY 138 RELQLRSLTELKGVLIQRFQCYQDTILKWDIFHKNQLALTLDTNRSRACHPCSP 197
DB 204 YTLIPDLRLVNGQVGFHNLYNLCNMTIQWSEIVSNGTDAYNYDFTAPERECPKCHE 263
QY 198 MCKSGRCWGSSEDCQLSTRVTCAGGCA--RCKGPLEPTDCHECAAGCTGPKHSDCLAC 255
DB 264 SCTHG-CWGECPKNCQKXSLTCSPPCAGGRCYCPKRECH-PCAGGCTGTQKDCIAC 322
QY 256 LHNHSGICELHCPALVYNTDTFESMNPBGRYTFGASCVTACPNYLSLTDVGSCTLYC 315
DB 323 KNFFDEAVSKEECPMKRYKAPTYVLETNPEGRYATGATCKECP-GHLRLDNGACVRSC 381
QY 316 PLHNQEVTAEDGTQRCCKSKPCAR 340
DB 382 PQDRMDKGGE-----CVPNGCPCK 401

RESULT 13
LT23 CABEL STANDARD; PRT; 1367 AA.
AC P24348;
DT 01-MAR-1992 (Rel. 21, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Let-23 receptor protein-tyrosine kinase precursor (EC 2.7.1.12).
OS Let-23 OR KIN-7 OR ZK1067.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.
RX NCBI_TaxID=6239;
RP SEQUENCE FROM N.A.
RX MEDLINE=91080919; PubMed=1979659;
RA Aroian R.V., Koga M., Mendel J.E., Ohshima Y., Sternberg P.W.;
RT "The let-23 gene necessary for Caenorhabditis elegans vulval induction encodes a tyrosine kinase of the EGF receptor subfamily.";
RL Nature 348:693-693(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=96177760; PubMed=8604137;
RA Sakai T., Koga M., Ohshima Y.;
RT "Genomic structure and 5' regulatory regions of the let-23 gene in the nematode C. elegans.";
RL J. Mol. Biol. 256:548-555(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Thomas K.;
RT Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP REVISIONS.
RA Durbin R.;
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP MUTANTS.
RX MEDLINE=94147981; PubMed=8313880;
RA Aroian R.V., Les G.M., Sternberg P.W.;
RT "Mutations in the Caenorhabditis elegans let-23 EGFR-like gene define elements important for cell-type specificity and function.";
RL EMBO J. 13:360-366(1994).
RN [6]
RP SUBCELLULAR LOCATION.
RC STRAIN=Bristol N2;
RX MEDLINE=99287744; PubMed=10359617;
RA Whittfield C.W., Barnes T., Hekimi S., Kim S.K.;
RT "Basolateral localization of the Caenorhabditis elegans epidermal growth factor receptor in epithelial cells by the PDZ protein lin-10.";
RL Mol. Biol. Cell 10:2087-2100(1999).
CC -!- FUNCTION: Tyrosine kinase receptor required for the induction of vulval differentiation. Possible receptor for the inductive signal required for vulval development. Activated by lin-3 and acts by way of let-60 RAS. The lin-3/let-23 pair is a simplified version of the mammalian neuroligin-ERBB network.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Basolateral and apical membrane of cell junctions in epithelial vulval precursor cells.
CC -!- TISSUE SPECIFICITY: Vulval precursor cells.
CC -!- DEVELOPMENTAL STAGE: Expressed during L2 and L3 larval stages.
CC -!- SIMILARITY: Belongs to the EGF receptor family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X57767; CAA40919.1; -
DR EMBL; D63426; BAA09729.1; -
DR EMBL; Z70038; CAA93882.2; -
DR PIR; S70712; S70712.
DR HSP; P11362; IFGK.
DR WormPep; ZK1067.1; CE25678.
DR InterPro; IPR000494; EGFR_L domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR009030; Grow_fac_recep.

```
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR008266; Tyr_kinase_AS.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF00069; kinase; 1.
DR Pfam: PF01030; Recep_L domain; 2.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: P0000001; Prot_kinase; 1.
DR SMART: SMO0220; S_TK; 1.
DR SMART: SMO0219; TyRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferrase; Tyrosine-protein kinase; Developmental protein;
KW ATP-binding; Phosphorylation; Transmembrane; Glycoprotein; Receptor;
KW Signal.
FT CHAIN 1 20
FT DOMAIN 21 1367
FT TRANSMEM 863 883
FT DOMAIN 884 1367
FT NP_BIND 935 943
FT BINDING 963 963
FT ACT_SITE 1054 1054
FT DISULFID 264 272
FT DISULFID 268 280
FT DISULFID 288 295
FT DISULFID 292 306
FT DISULFID 307 315
FT DISULFID 311 323
FT DISULFID 326 335
FT DISULFID 339 366
FT DISULFID 370 381
FT DISULFID 385 400
FT DISULFID 403 408
FT DISULFID 584 573
FT DISULFID 568 581
FT DISULFID 584 593
FT DISULFID 597 611
FT DISULFID 614 621
FT DISULFID 618 629
FT DISULFID 632 648
FT DISULFID 652 664
FT DISULFID 667 676
FT DISULFID 671 688
FT DISULFID 691 704
FT DISULFID 714 737
FT DISULFID 740 747
FT DISULFID 744 759
FT DISULFID 761 775
FT DISULFID 779 794
FT DISULFID 797 807
FT DISULFID 801 815
FT DISULFID 818 831
FT DISULFID 835 849
FT CARBOHYD 135 135
FT CARBOHYD 213 213
FT CARBOHYD 229 239
FT CARBOHYD 420 420
FT CARBOHYD 605 605
FT CARBOHYD 699 699
FT CARBOHYD 790 790
FT CARBOHYD 820 820
FT VARIANT 412 412
FT VARIANT 513 513
FT VARIANT 744 744
FT VARIANT 797 797
FT VARIANT 1109 1109
FT VARIANT 1118 1118
FT CONFLICT 44 44
SEQUENCE 1367 AA; 155561 MW; A46239431DC02E33 CRC64;

Query Match 18.3%; Score 419; DB 1; Length 1367;
Best Local Similarity 29.0%; Pred. No. 1.le-24;
Matches 102; Conservative 59; Mismatches 131; Indels 60; Gaps 14;

QY 25 VCTGDKMLRLPASPETHLDMLRLHYQCVQVQNLLELYLPTN----- 68
DQ 83 LCGTNGISRYGNT-LEDLETMYGCRVYVGNLEITWIEANEIKKKESTNSTVDPK 141
QY 69 -----ASLSFLQDIQEVQGVVLIHAHQVQVQLRRLIVRGVQTLQFENYALAVLDNGDP 122
DQ 142 NEDSPLKSIINFFDNLLEIRGSLIIYRANIKISFRLAVIYGVDEVDHN-ALYHKNDK- 199
QY 123 LNNTPVTGASPGGLRELQLRLSLTEILKGGVLIQRPQLCY-QDTILWKDIFHKNNQAL 181
DQ 200 -----VHEVWRELVRVIRNGSVTIQDNFMKCYIGSKIDWKELLYDPD--VQ 243
QY 182 TLITNRSRACH-----PCSPMKGSRGSESEDCOSLTRVTCAGGCARC---KGPL 231
DQ 244 KVETTNHSHQYQNGKSNMAKCHESC-NDKMGSGDNDQQRVYRVCPRSCSCQCFVSNST 302
QY 232 PTDCHQEQCAAGCTGPKHSDCLACLFHFNHSHGICELHCPALVTYNTDTFESMNPPEGRTY 291
DQ 303 SYECDSACLGGCTGHHGPKNCIACSKYELDGICETCTPSRKIFNHKIGRLVFNPDGRYON 362
QY 292 GASCVTACPNYL-STDVGSCTLYC-PLHNOEVTAEQDQCEKC-SKPCAR 340
DQ 363 GNHCVKCPPELLIENDV--CVRECSGDGHYATKD--VRECEKCRSSSCPK 410

RESULT 14
ILPR_BRALA STANDARD; PRT; 1363 AA.
AC 002466;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Insulin-like peptide receptor precursor (EC 2.7.1.112) (ILP receptor).
OS Branchiostoma lanceolatum (Common lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
CC NCBI TaxID=7740;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96408719; PubMed=9813726;
RA Fashmforoush W., Chan S.J., Steiner D.F.;
RT "Structure and expression of the insulin-like peptide receptor from
  amphioxus.";
RL Mol. Endocrinol. 10:857-866(1996).
CC -|- FUNCTION: THIS RECEPTOR BINDS TO THE INSULIN RELATED PEPTIDE AND
  HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC -|- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
  tyrosine phosphate.
CC -|- SUBUNIT: PROBABLE TETRAMER OF 2 ALPHA AND 2 BETA CHAINS LINKED BY
  DISULFIDE BONDS. THE ALPHA CHAINS CONTRIBUTE TO THE FORMATION OF
  THE LIGAND-BINDING DOMAIN, WHILE THE BETA CHAIN CARRY THE KINASE
  DOMAIN (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -|- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
  receptor subfamily.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
EMBL; S83394; AAB50848.1; -.
PIR; T43220; T43220.
HSSP; P06213; iIRK.
InterPro; IPR000494; EGFR_L_domain.
```

```

DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow_fac_recep.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002011; RecepttyrkinsII.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF00041; fn3; 3.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SMO0060; FN3; 3.
DR SMART; SMO0261; FU; 1.
DR SMART; SMO0219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Glycoprotein; ATP-binding; Phosphorylation; Signal.
DR SIGNAL 1 29
DR CHAIN 30 716
FT PROPEP 717 720
FT CHAIN 721 1363
FT DOMAIN 721 928
FT TRANSFEM 929 949
FT DOMAIN 950 1363
FT NP_BIND 1000 1008
FT BINDING 1028 1028
FT ACT_SITE 1148 1148
FT MOD_RES 1174 1174
FT CARBOHYD 51 51
FT CARBOHYD 97 97
FT CARBOHYD 137 137
FT CARBOHYD 278 278
FT CARBOHYD 483 483
FT CARBOHYD 599 599
FT CARBOHYD 617 617
FT CARBOHYD 665 665
FT CARBOHYD 666 666
FT CARBOHYD 711 711
FT CARBOHYD 732 732
FT CARBOHYD 736 736
FT CARBOHYD 743 743
FT CARBOHYD 816 816
FT CARBOHYD 885 885
FT CARBOHYD 898 898
FT SEQUENCE 1363 AA; 154104 MW; 238120B4EAB1ED65 CRC64;
Query Match 15.0%; Score 342.5; DB 1; Length 1363;
Best Local Similarity 28.3%; Pred. No. 9.3e-19;
Matches 106; Conservative 40; Mismatches 124; Indels 105; Gaps 20;
Qy 9 WGLL-----LALLPFGAASQTGCTDMKRLPASPETHLDMLRHLYGQCGWQGNLELT 63
Db 12 WAULTLVIGLLVPSNGEEVYCSDMRNK-----VSNLQQL-ENCTVIEGYLQI- 61
Qy 64 YLPNTASLSFLDQCEVQGVLLIAHQVRQP-----LQLRL-----IVRGTO 106
Db 62 -----LLIDFAEEQDSGLAFPNLVITDVFLLYVRGLTNLSLSEFNIAVINGTN 112
Qy 1.07 LFDNYALAVLDNGDPLNNTTPVTGASPGGLREQLRSLTEILKGGVLIORNPOLCYQDT 166
Db 113 LF-FNYALVPEMLD-----MQKIGLYSLQNIITGSRVIEKNPNLCYLDLT 156
Qy 167 ILWKDIF---HNNQLALTLIDTVRSRAC-HPQSPMCK-----GSRCKGSESSDCQSLT 216

```

```

Db 157 IDMSFIAESGYNN-----FIVNREBEHCNVFCPCRIKHVPQLDLQWAE--EHCKVC 210
Qy 217 RTVCAGGCARCKGPIPLTDCHEQCAAGCTGPKHSDCLACHFNHSGICELHCPALVTYNT 276
Db 211 PSSCLNCR-----DGISGCCHEHCIGGCDGPTERDCVACKYFVHNGECLIQCPPTQYXK 266
Qy 277 D-----TFESMNPBGRY--TFGASCVTACPVNVLSTDVG-----CTLVCPLNQEVTAEDG 327
Db 267 DRCITEECPTNTNSVKLHKRKCIPCEPSGY-TIDINNPLCT----- 310
Qy 328 TORCE-KCSKPCARG 341
Db 311 --ECEGQCPKCKGG 323
RESULT 15
INSTR DROME
ID INSTR DROME STANDARD; PRT; 2146 AA.
AC P09208; Q24023; Q24089;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE insulin-like receptor precursor (EC 2.7.1.112).
GN INR OR INR-A OR DIR-A.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95354655; PubMed=7628438;
RA Fernandez R., Tabarini D., Azpiazu N., Frasch M., Schlessinger J.;
RT "The Drosophila insulin receptor homolog: a gene essential for
embryonic development encodes two receptor isoforms with different
signaling potential.";
RL ENBO J. 14:3373-3384(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95181404; PubMed=7876183;
RA Ruan Y., Chen C., Cao Y., Garofalo R.S.;
RT "The Drosophila insulin receptor contains a novel carboxyl-terminal
extension likely to play an important role in signal transduction.";
RL J. Biol. Chem. 270:4236-4243(1995).
RN [3]
RP SEQUENCE OF 652-1749 FROM N.A.
RX STRAIN=Oregon-R; TISSUE=Embryo;
RX MEDLINE=87100165; PubMed=3099787;
RA Nishida Y., Hata M., Nishizuka Y., Rutter W.J., Ebina Y.;
RT "Cloning of a Drosophila cDNA encoding a polypeptide similar to the
human insulin receptor precursor.";
RL Biochem. Biophys. Res. Commun. 141:474-481(1986).
RN [4]
RP SEQUENCE OF 1297-1595 FROM N.A.
RX MEDLINE=86259667; PubMed=3014506;
RA Petruzzelli L., Herrera R., Arenas-Garcia R., Fernandez R.,
Birnbbaum M.J., Rosen O.M.;
RT "Isolation of a Drosophila genomic sequence homologous to the kinase
domain of the human insulin receptor and detection of the
phosphorylated Drosophila receptor with an anti-peptide antibody.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:4710-4714(1986).
CC -1- FUNCTION: THIS RECEPTOR PROBABLY BINDS AN INSULIN RELATED PROTEIN
AND HAS A TYROSINE-PROTEIN KINASE ACTIVITY (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
CC -1- SUBUNIT: Tetramer of 2 alpha and 2 beta chains linked by disulfide
bonds. The alpha chains contribute to the formation of the ligand-
binding domain, while the beta chains carry the kinase domain.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
receptor subfamily.

```



```

QY 331 CE-----KSCPKCARGTHSLPLR 348
DB 635 CEPKCGKCDKRCSSGLIDSLR 657

RESULT 16
HTK7 HYDAT
ID HTK7 HYDAT STANDARD; PRT; 1477 AA.
AC Q25197;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative insulin-like peptide receptor precursor (EC 2.7.1.112).
GN HTK7.
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroidea; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN SEQUENCE FROM N.A.
RP Steele R.E., Mai N.H., Lieu P., Shenk M.A.;
RA Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.
RL FUNCTION: THIS RECEPTOR PROBABLY BINDS AN INSULIN RELATED PROTEIN
CC -!- AND HAS A TYROSINE-PROTEIN KINASE ACTIVITY (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN DIVIDING EPITHELIAL CELLS.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
CC receptor subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.iesb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; M64612; AAA68205.1;
CC PIR; T18534; T18534.
CC HSSP; P06213; IIRK.
CC InterPro; IPR000494; EGFR_L domain.
CC InterPro; IPR003961; FN III.
CC InterPro; IPR006211; Furin-like.
CC InterPro; IPR008212; Furin repeat.
CC InterPro; IPR009030; Grow_fac_recep.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002011; RecepttyrkinsII.
CC InterPro; IPR001245; Tyr_pkinase.
CC InterPro; IPR008266; Tyr_pkinase_AS.
CC Pfam; PF00041; fn3; 1.
CC Pfam; PF00757; Furin-like; 1.
CC Pfam; PF00069; Pkinase; 1.
CC Pfam; PF01030; Recep_L domain; 2.
CC PRINTS; PR0109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00060; FN3; 2.
CC SMART; SM00261; FU; 1.
CC SMART; SM00219; Tyr_KC; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
CC PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
CC Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;
CC Glycoprotein; ATP-binding; Phosphorylation; Signal.
FT SIGNAL 1-24 POTENTIAL.
FT CHAIN 25 1477 PUTATIVE INSULIN-LIKE PEPTIDE RECEPTOR.
FT DOMAIN 25 980 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 981 1001 POTENTIAL.
FT DOMAIN 1002 1477 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1044 1315 PROTEIN KINASE.

```

```

PT NP BIND 1050 1058 ATP (BY SIMILARITY).
PT BINDING 1077 1077 ATP (BY SIMILARITY).
PT ACT SITE 1175 1175 BY SIMILARITY.
PT MOD_RES 1201 1201 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 55 55 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 325 325 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 491 491 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 549 549 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 644 644 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 732 732 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 791 791 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 874 874 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 895 895 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 957 957 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 1477 AA; 168276 MW; 74ACDBA7C6DE1D41 CRC64;

Query Match 12.7%; Score 291; DB 1; Length 1477;
Best Local Similarity 29.2%; Pred. No. 1e-14;
Matches 94; Conservative 34; Mismatches 114; Indels 80; Gaps 18;

QY 51 QCGVVOGNI---ELTYLPTNASLSFLQDIQVGVLIHNRQVPL-QRLRIVRGTC 106
DB 54 QNCTCHGNDLVKSTKYDEENFKYFPKRLITGHLISLCTLKFFHPPGLTVIRGGD 113
QY 107 LFDNYVALVLDNGDFLNNTPVTGASPGGLRELRLSLTEILKGVLIQRPOLCYQDT 166
DB 114 LIL-NYALVIYN-----EIKEVVPPSLTAILNGGVHIGRHLRCYVNT 156
QY 167 ILWKDI---PHKNNQLALTLDNRSRA-----CHPC-----SPMKGSRGWG--- 206
DB 157 IRWSIKIHQGGQGIYL-ESNKLNDLGLKCHGHPAGHDGP--KAQYCWGPGPK 213
QY 207 --ESSDCDSLTRTCAGGACRCKGELPTDCCHQCACAGCTGPKHSD-CLACLHF--NHS 261
DB 214 KQNKACQACQCFNCTQC--GPEGCLDGSDHICCHHECLGCSAINSTNTCHACKRYIKST 271
QY 262 GICELHCPALVNTVTF---ESMP-----NPEGRYTFGASCTVACPNNYLSLTDVGSCT 312
DB 272 GQCVSKQPR-KQYLVKFLQCESCPWNSINSTEYHYLWQGCVCYKPNYIS----- 323
QY 313 LVCPLEHNEVTAADGTQRCCK 334
DB 324 -----NNQ-----TKCKCK 333

RESULT 17
IRR MOUSE
ID IRR MOUSE STANDARD; PRT; 1300 AA.
AC Q9WTL4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Insulin receptor-related protein precursor (EC 2.7.1.112) (IRR)
DE (IR-related receptor).
GN INSR OR IRR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=99272338; PubMed=10342810;
RA Hirayama I., Tamemoto H., Yokota H., Kubo S.-K., Wang J., Kuwano H.,
RA Nagamachi Y., Takeuchi T., Izumi T.;
RT "Insulin receptor-related receptor is expressed in pancreatic b-cells
RT and stimulates tyrosine phosphorylation of insulin receptor
RT substrate-1 and-2."
RL Diabetes 48:1237-1244 (1999).
CC -!- FUNCTION: This receptor probably binds an insulin related protein
CC and has a tyrosine-protein kinase activity. It phosphorylates the

```

insulin receptor substrates IRS-1 and IRS-2.
 -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 tyrosine phosphate.
 -!- SUBUNIT: PROBABLE Tetramer of 2 alpha and 2 beta chains linked by
 disulfide bonds. The alpha chains contribute to the formation of
 the ligand-binding domain, while the beta chains carry the kinase
 domain (by similarity).
 -!- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
 -!- TISSUE SPECIFICITY: Highly expressed in the islets as well as in
 pancreatic beta-cells.
 -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
 receptor subfamily.
 -!- SIMILARITY: Contains 3 fibronectin type III domains.
 This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; AB007135; BAA77835.1; --
 DR HSP; P06213; IIRK.
 DR MGD; MGI:1346037; Inerr.
 DR InterPro; IPR000494; EGFR_L_domain.
 DR InterPro; IPR003957; FN_III-like.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR006211; Furin-like.
 DR InterPro; IPR006212; Furin repeat.
 DR InterPro; IPR009030; Grow fac recep.
 DR InterPro; IPR009719; Prot_kinase.
 DR InterPro; IPR002011; RecepttyrkinasII.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR InterPro; IPR008266; Tyr_pkinase_AS.
 DR Pfam; PF00041; fn3; 2.
 DR Pfam; PF00757; Furin-like; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF01030; Recep_L_domain; 2.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00060; FN3; 3.
 DR SMART; SM00261; FU; 1.
 DR SMART; SM00219; TyrK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
 KW Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
 Glycoprotein; ATP-binding; Phosphorylation; Repeat; Signal.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 1300 INSULIN RECEPTOR-RELATED PROTEIN, ALPHA-
 CHAIN (PROBABLE).
 FT CHAIN 27 746 INSULIN RECEPTOR-RELATED PROTEIN, BETA-
 CHAIN (PROBABLE).
 FT CHAIN 747 1300
 FT CHAIN 747 921
 FT DOMAIN 747 921
 FT TRANSMEM 922 943
 FT DOMAIN 944 1300
 FT DOMAIN 483 601
 FT DOMAIN 602 811
 FT DOMAIN 813 910
 FT DOMAIN 979 1254
 FT NP_BIND 985 993
 FT BINDING 1013 1013
 FT ACT_SITE 1115 1115
 FT DISULFID 214 222
 FT DISULFID 216 228
 FT DISULFID 229 237
 FT DISULFID 233 246
 FT DISULFID 249 258
 FT DISULFID 262 274
 FT DISULFID 260 300

FT DISULFID 304 317 BY SIMILARITY.
 FT DISULFID 320 324 BY SIMILARITY.
 FT DISULFID 657 864 INTERCHAIN (POTENTIAL).
 FT CARBOHYD 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 311 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 411 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 492 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 528 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 516 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 534 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 756 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 885 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 898 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT MOD_RES 1145 1145 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 SQ SEQUENCE 1300 AA; 144745 MW; 2989FC3521A2067 CRC64;
 Query Match 12.7%; Score 290; DB 1; Length 1300;
 Best Local Similarity 28.1%; Pred. No. 1.1e-14;
 Matches 112; Conservative 52; Mismatches 118; Indels 116; Gaps 26;
 QY 1 MELAALCRWG--LILALLPPCAA--STQVCTGTDMKRLRPASPETHLDMLHLYQGQVW 56
 DB 1 MAVPALNFWGVHLLMSLLSGLDTEVCPSLDIR-----SEVTELRL-ENCSVV 51
 QY 57 QGNLE--LTYLPTN---ASLSFLQDIOEVGVLIHNNQVQVPLQRLR-----IVRG 104
 DB 52 EGHLLQILMFANAGEDPRGLSFFR-LTQVTDYLL-----FRVYGLSLRDLFPNLTVIRG 106
 QY 105 TQLFEDNYVALVLDNGDPLNNTTPTVTCASPGGLRELQRLSLTEILKGVLQIORNPOLCYQ 164
 DB 107 TRLEPL-GYALLIEMPH-----LADVGLPSLGAIVRGAVRKNQELCHL 150
 QY 165 DTILWMDIFHNNQLALTIDTNR---SRACHPCSPMKCG-----S 202
 DB 151 STIDW-----GLLPAPGTNHNIGKLGECADVCVGLGAAGFCSTRTTFSGRDY 202
 QY 203 RCGESSEDCOSLTRTYCAGGCARCKGLPTDCCHCOCAAGCTGPKH-SDCLACLFHNS 261
 DB 203 RCW--TSSHCQKV--CPCPRGMACTAG---GDCCHSELGCGSQFEDPRACVACHRYFQ 255
 QY 262 GICELHCPALVTYNTDTFESMPNPEGRYTFEGASCVTA--CPY-----NYLSTDV---GSC 311
 DB 256 GVCRLACPP---GTQYES-----GR-----CVTAELCAHLREVPGLATTPIGYEGSC 300
 QY 312 TLVCPHLNQVTAEDGTQCEKCS---KPCARGTHSL 345
 DB 301 LAQCP---PGFTRNGSSIFCKCEGLCPCKEKGVTGKITI 335
 RESULT 18
 ID - IRR_CAVPO STANDARD; PRT; 1300 AA.
 AC P14617;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Insulin receptor-related protein precursor (EC 2.7.1.112) (IRR)
 DE (IR-related receptor).
 GN INSR.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89359245; PubMed=2768234;
 RA Shier P., Watt V.M.;
 RT "Primary structure of a putative receptor for a ligand of the insulin
 family.";
 RL J. Biol. Chem. 264:14605-14608(1989).
 CC -!- FUNCTION: This receptor probably binds an insulin related protein
 and has a tyrosine-protein kinase activity. It phosphorylates the

CC -!- FUNCTION: This receptor probably binds an insulin-related protein and has a tyrosine-protein kinase activity. It phosphorylates the insulin receptor substrates IRS-1 and IRS-2.

CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.

CC -!- SUBUNIT: PROBABLE Tetramer of 2 alpha and 2 beta chains linked by disulfide bonds. The alpha chains contribute to the formation of the ligand-binding domain, while the beta chain carry the kinase domain (by similarity).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).

CC -!- TISSUE SPECIFICITY: EXPRESSED PREFERENTIALLY IN THE KIDNEY. ALSO FOUND IN STOMACH AND THYMUS BUT NOT IN SKELETAL MUSCLE, BRAIN, INTESTINE, AND UTERUS.

CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin receptor subfamily.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL: M90661; AAB59692.1; -

CC DR EMBL: M90660; AAA41452.1; -

CC EMBL: D12678; BAA20982.1; -

CC PIR: PC1130; PC1130.

CC InterPro: IPR006212; Furin repeat.

CC InterPro: IPR009030; Grom_Fac_recep.

CC InterPro: IPR000719; Prot_kinase.

CC InterPro: IPR02011; RecepttyrkinasII.

CC InterPro: IPR008266; Tyr_kinase_AS.

CC SMART: SMO0261; FU; 1.

CC PROSITE: PS00107; PROTEIN KINASE ATP; PARTIAL.

CC PROSITE: PS00109; PROTEIN KINASE TYR; PARTIAL.

CC PROSITE: PS00011; PROTEIN KINASE DOM; PARTIAL.

CC PROSITE: PS00239; RECEPTOR_TYR_KIN_II; PARTIAL.

CC Transferrase: Tyrosine-protein kinase; Receptor; Glycoprotein;

CC KW ATP-binding; Phosphorylation; Signal.

CC FT SIGNAL 1 26 POTENTIAL.

CC FT CHAIN 27 >581 INSULIN RECEPTOR-RELATED PROTEIN.

CC FT DISULFID 214 222 BY SIMILARITY.

CC FT DISULFID 216 228 BY SIMILARITY.

CC FT DISULFID 229 237 BY SIMILARITY.

CC FT DISULFID 233 246 BY SIMILARITY.

CC FT DISULFID 249 258 BY SIMILARITY.

CC FT DISULFID 262 274 BY SIMILARITY.

CC FT DISULFID 280 300 BY SIMILARITY.

CC FT DISULFID 304 317 BY SIMILARITY.

CC FT DISULFID 320 324 BY SIMILARITY.

CC FT CARBOHYD 47 47 N-LINKED (GLCNAC. .) (POTENTIAL).

CC FT CARBOHYD 311 311 N-LINKED (GLCNAC. .) (POTENTIAL).

CC FT CARBOHYD 411 411 N-LINKED (GLCNAC. .) (POTENTIAL).

CC FT NON_CONS 481 482 N-LINKED (GLCNAC. .) (POTENTIAL).

CC FT CARBOHYD 483 483 N-LINKED (GLCNAC. .) (POTENTIAL).

CC FT CARBOHYD 534 534 N-LINKED (GLCNAC. .) (POTENTIAL).

CC FT VARIANT 425 425 M -> L.

CC FT NON_TER 581 581

CC SQ SEQUENCE 581 AA; 430PA6E1498C3BB9 CRC64;

Query Match 11.8%; Score 270; DB 1; Length 581;

Best Local Similarity 26.4%; Pred. No. 1.5e-13;

Matches 104; Conservative 59; Mismatches 123; Indels 108; Gaps 24;

QY 1 MELAALCRWG-LILLALLPQCA--STQVCTGTDMLRLPASPTHLMDLRLHLYQGQGVV 56

DB 1 MAYPALWPGVGYLLMSLLSLGSCLDTUEVCPSLDI-----SEVTELRRL-ENGSVV 51

QY 57 QGNLE-LTYLPTN---ASLSFLQDIOEGVGYVLIANQVRQVFLQRL-----IVRG 104

DB 52 EGHLOILMFAATGEDERGSFPR-LTQVTDYLL-----FRVYGLSLRLFPNLAIVRG 105

```

QY 105 TOLPEDYVALVLDNGDFPLNNPTVTVTGASPGGLRELQLRSLTBLIKGGVLIOENPOLCYQ 164
DB 107 ARULFL-GYALIIIFEMPH-----LRDGLPSLGLVAGAVVERKQELCHL 150
QY 165 DTILWKDIFHKNNQLALTIDITNRSRACHPCSPMKCG-----SRCW 205
DB 151 STIDWGLL-----QPAPGANHIYGNKLBECADVCPGVLAGAGEPCVRTTFGGHTIDYRCW 205
QY 206 GESSEPCOSLTRVCAGGACRKGPLPTDCCHEOCAAGCTGPKH-SDCLACLHFNHSGIC 264
DB 206 --TSSHCQRV--CPCPRGLACTVG--GECCHSECLGGCSQPDPRACVACRHLIFQGVC 258
QY 265 ELHCPALVTYNTDTFESMPN-----PEGRYTFG---ASCVTACPNYSLTDVGSCT 312
DB 259 LPACPP-GTYQYESSWCRTAELGHLREVPGHATAFGIYEGSCLAQCPGCF--TRNGS-S 314
QY 313 LVPELHNQEVTAEDGTCQCE-KCSKFCARGTHSL 345
DB 315 IFC-----HKCEGLCPKECKVGTKTI 335

RESULT 20
MIPR LYMST STANDARD; PRT; 1607 AA.
ID MIPR LYMST Q25410;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative molluscan insulin-related peptide(S) receptor precursor
   (EC 2.7.1.112).
OS Lymnaea stagnalis (Great pond snail).
OC Bukaryaota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidea; Lymnaeidae; Lymnaea.
OX NCBI_TaxID=6523;
EN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CNS;
RX MEDLINE=96032341; PubMed=7557427;
RA Roovers E., Vincent M., van Kesteren E., Geraerts W.P.M.,
RA Planta R.J., Vreugdenhil E., van Heerikhuizen H.;
RT "Characterization of a putative molluscan insulin-related peptide
   receptor.";
RL Gene 162:181-198(1995).
CC -!: FUNCTION: THIS RECEPTOR PROBABLY BINDS TO THE FOUR DIFFERENT
CC MOLLUSCAN INSULIN-RELATED PEPTIDES AND HAS A TYROSINE-PROTEIN
CC KINASE ACTIVITY.
CC -!: CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!: SUBUNIT: PROBABLE TETRAMER OF 2 ALPHA AND 2 BETA CHAINS LINKED BY
CC DISULFIDE BONDS. THE ALPHA CHAINS CONTRIBUTE TO THE FORMATION OF
CC THE LIGAND-BINDING DOMAIN, WHILE THE BETA CHAIN CARRY THE KINASE
CC DOMAIN (BY SIMILARITY).
CC -!: SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!: SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
CC receptor subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X84994; CAA59353.1; -.
CC PIR; T43212; T43212.
CC HSP; P06213; IIRK.
CC InterPro; IPR000494; EGFR_L domain.
CC InterPro; IPR003961; FN_III-
CC InterPro; IPR062211; Furin-like.
CC InterPro; IPR006212; Furin repeat.
CC InterPro; IPR009030; Growth fac recep.
CC InterPro; IPR000719; Prot kinase.

```

```

DR InterPro: IPR002011; RecepttyrkinasII.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR00266; Tyr_kinase_AS.
DR Pfam: PF00041; fr3.1.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF01030; Recep_L domain; 2.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PDC00001; Prot_Kinase; 1.
DR SMART: SMO0060; FN3; 3.
DR SMART: SMO0261; FU; 1.
DR SMART: SMO0219; TyrKc; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS00109; PROTEIN KINASE_TYR; 1.
DR PROSITE: PS00239; RECEPTOR_TYR_KIN_II; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR Kx Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;
Kw Glycoprotein; ATP-binding; Phosphorylation; Signal.
FT SIGNAL 1 35
FT CHAIN 36 694
FT PROPEP 695 697
FT CHAIN 698 1607
FT DOMAIN 698 975
FT TRANSMEM 976 996
FT DOMAIN 997 1607
FT DOMAIN 1017 1308
FT NP_BIND 1043 1051
FT BINDING 1072 1072
FT ACT_SITE 1173 1173
FT MOD_RES 1199 1199
FT CARBOHYD 82 82
FT CARBOHYD 188 188
FT CARBOHYD 245 245
FT CARBOHYD 275 275
FT CARBOHYD 332 332
FT CARBOHYD 343 343
FT CARBOHYD 495 495
FT CARBOHYD 520 520
FT CARBOHYD 563 563
FT CARBOHYD 710 710
FT CARBOHYD 778 778
FT CARBOHYD 796 796
FT CARBOHYD 802 802
FT CARBOHYD 868 868
FT CARBOHYD 879 879
FT CARBOHYD 940 940
FT CARBOHYD 953 953
FT SEQUENCE 1607 AA; 181820 MW; P9CDIAE325D2ED6 CRC64;
Query Match 11.8%; Score 269.5; DB 1; Length 1607;
Best Local Similarity 23.1%; Pred. No. 5.4e-13;
Matches 88; Conservative 57; Mismatches 131; Indels 105; Gaps 17;
QY 25 VCTGTDMLRLPASPETHLDMLRHLVQGVVQGNLELTPLTNA---SLSLQDIQEV 80
DB 64 VCGSVDIR-----SSMDNFK-ILENCWTIEGSLRISLPKALDFPHSLP-PDLREI 113
QY 81 QGYVLIANQNRQVPLQ-----RLRIVRGQLPFEDNYALVDNGPLNNTFTVTGAS 133
DB 114 TDYLLM---YRVYGLTSLKLPFNLAIRGSELF-NSYAIWVEMRD----- 156
QY 134 PGLRELQRLSRTEILKGVLTQRPOLCYQDTILWKKDIFHNKNQLALTLDITNRSRACH 193
DB 157 ---LQDLGLVNLRTISRGVRLTKRFLKCYIETINWTOIGVSDPARRFI--NNKEQCPN 211
QY 194 PCSPMCKGRWCW----ESSEDCQSLTRTVTCAGGCARCKGKPLPTDCCHCQAAGCTGPKH 249
DB 212 SKCDEQSRWCWYSDYSCQKGLNCQCKENTYCMEN-----GSCCHDYCLGGCKVPMN 262
QY 250 SD-CLACLHFNHSGICELHCP-----ALVTYNTDTTFESMENPEGRYTFG 292

```

263 PDECFCKEYQFNNTCEPQCPCPTTYKFLNRCCLITDKECLALTNDPGNTPKLLDGEKGP 322

293 ASCVTACPNYSLTDVG-----SCLVCP--LHNQVTAEDGTQCEKSKPC 338

323 SLCLYTCPNY-----SVGSDKNKNSQCVKRCQLCPKSGHLEINNIQDAHKLKESK-- 377

339 ARGTHSLPRPAAVPPPLRMQ 359

378 -----ISGFLXIQ 385

RESULT 21

IRR_HUMAN

ID IRR_HUMAN STANDARD; PRT; 1297 AA.

AC P14616; O60724;

DT 01-APR-1990 (Rel. 14, Created)

DT 18-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Insulin receptor-related protein precursor (EC 2.7.1.112) (IRR)

DE (IR-related receptor).

DE INSR OR IRR.

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RC MEDLINE=99243372; PubMed=10226785;

RA Haenze J., Berthold A., Klammt J., Gallaher B., Siebler T.,

RA Kratsch J., Emlinger M., Kiess W.,

RA "Cloning and sequencing of the complete cDNA encoding the human

RI Insulin receptor related receptor.";

RI Horm. Metab. Res. 31:77-79(1999).

[2]

SEQUENCE OF 30-1297 FROM N.A.

MEDLINE=89359245; PubMed=2768234;

RA Shier P., Watt V.M.;

RI "Primary structure of a putative receptor for a ligand of the insulin

RI family.";

RI J. Biol. Chem. 264:14605-14608(1989).

CC -!- FUNCTION: This receptor probably binds an insulin related protein

and has a tyrosine-protein kinase activity. It phosphorylates the

insulin receptor substrates IRS-1 and IRS-2.

CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein

tyrosine phosphate.

CC -!- SUBUNIT: PROBABLE TETRAMER OF 2 ALPHA AND 2 BETA CHAINS LINKED BY

DISULFIDE BONDS. THE ALPHA CHAINS CONTRIBUTE TO THE FORMATION OF

THE LIGAND-BINDING DOMAIN, WHILE THE BETA CHAIN CARRY THE KINASE

DOMAIN.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin

receptor subfamily.

CC -!- SIMILARITY: Contains 3 fibronectin type III domains.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@sib-sib.ch).

CC

CC EMBL; AF064078; AAC17167.1; -

CC EMBL; J05046; AAC31759.1; -

CC HSSP; P04213; IIRK

CC Genew; HGNC:6093; INSR.

CC MIM; 147671; -

CC GO; GO:0005887; C: integral to plasma membrane; NAS.

CC GO; GO:0004714; F: transmembrane receptor protein tyrosine kin. ; NAS.

CC GO; GO:0007169; P: transmembrane receptor protein tyrosine kin. ; NAS.

CC InterPro; IPR000494; EGFR_L domain.

CC InterPro; IPR008957; FN_III-like.

```
DR InterPro: IPR003961; FN III.
DR InterPro: IPR008211; Furin-like.
DR InterPro: IPR008212; Furin-repeat.
DR InterPro: IPR009030; Grow fac recep.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002011; Recept tyrosine.
DR InterPro: IPR001245; Tyr kinase.
DR InterPro: IPR008266; Tyr_kinase_AS.
DR Pfam: PF00041; FN3; 2; 1.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF00669; Recept kinase; 1.
DR Pfam: PF01030; Recept L domain; 2.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot kinase; 1.
DR SMART: SM00060; FN3; 3.
DR SMART: SM00261; Fur; 1.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE: PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE: PS00239; RECEPTOR TYR KIN II; 1.
DR Transferrase: Tyrosine-protein kinase; Receptor; Transmembrane;
KW Glycoprotein; ATP-binding; Phosphorylation; Repeat; Signal.
FT SIGNAL 1 26
FT CHAIN 27 1297
FT CHAIN 27 746
FT CHAIN 747 1297
FT CHAIN 747 921
FT DOMAIN 747 921
FT TRANSMEM 922 943
FT DOMAIN 944 1297
FT DOMAIN 944 601
FT DOMAIN 602 811
FT DOMAIN 813 910
FT DOMAIN 910 921
FT DOMAIN 921 933
FT NP_BIND 933 993
FT BINDING 1013 1013
FT ACT_SITE 1115 1115
FT DISULFID 214 222
FT DISULFID 219 228
FT DISULFID 229 237
FT DISULFID 243 246
FT DISULFID 249 258
FT DISULFID 262 274
FT DISULFID 280 300
FT DISULFID 304 317
FT DISULFID 320 324
FT DISULFID 657 864
FT CARBOHYD 47 47
FT CARBOHYD 311 311
FT CARBOHYD 411 411
FT CARBOHYD 492 492
FT CARBOHYD 528 528
FT CARBOHYD 616 616
FT CARBOHYD 634 634
FT CARBOHYD 756 756
FT CARBOHYD 885 885
FT CARBOHYD 898 898
FT MOD_RES 1145 1145
FT SEQUENCE 1297 AA; 143719 MW; BB22C7FF61E3B65D CRC64;
Query Match 11.8%; Score 269; DB 1; Length 1297;
Best Local Similarity 25.4%; Pred. No. 4.6e-13;
Matches 101; Conservative 60; Mismatches 122; Indels 114; Gaps 23;
QY 1 MELAALCWGLLLAL-----LPPGAASCTVCTGTMKLRKLPASPETHLMKLRHLVQGCQV 56
DB 1 NAVPSLWPGACLPVIFSLGFGDTEVCPSLDIR-----SEVAELRQL-ENCSV 51
QY 57 QGNLELYLPTNA-----SLSLGDIQIEVQGVLLAHNQVQVLPQLR-----IVRG 104
DB 52 ECHQLILMTATGEDFRLSPFR-LTQVTDYLL-----FRVGLSLRDLFPNLAIVRG 106
QY 105 TQLFEDNYALVLDNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQNPOLCYQ 164
DB 107 TRFLP-GYALVIFEMPH-----LRDVALPALGAVLRGAVRVEKNQELCHL 150
QY 165 DTILWKDIFHNQNLALTLDITNRSRACHPCSPCKG-----SRGW 205
DB 151 STIDWGLL-----QPAPGANHIVGNKLBECADVCPGVLAGAGEPCAKTTFSGHTDYRCW 205
QY 206 GESSEDCOSLTRTVCA---GGCARKCKGLPTDCCHEOCAGCTGPKH-SDCLACLHFNHS 261
DB 206 --TSSHCQ-----RVCPCHGACTARG-----ECHTECLGCSQPEDPRACVACHRLYFQ 255
QY 262 GICELHCPALVTYNTDTE-----SMNPBGR-YTFG---ASCVTACPYNYLSTDVG 309
DB 256 GACLVACFP-GTYQYESWRCVTAERCASLHSPGRASTFGIHQSGCLAQCPSGFTR---N 311
QY 310 SCTLVCPHLNQEVTAEDGTORCE-KSKPCARGTHSL 345
DB 312 SSSIFC-----HKCEGLCPCKECKVGTXTI 335
RESULT 22
INSR_HUMAN STANDARD; PRT; 1382 AA.
ID INSR_HUMAN
AC P06213;
DT 01-JAN-1988 (Rel. 06, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Insulin receptor precursor (BC 2.7.1.112) (IR) (CD220 antigen).
GN INSR.
OS Homo sapiens (Human).
OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=85176928; PubMed=2859121;
RA Esina Y., Ellis L., Jarnagin K., Edery M., Graf L., Clauser E.,
RA Ou J.-H., Maslarsz F., Kan Y.W., Goldfine I.D., Roth R.A., Rutter W.J.;
RT "The human insulin receptor cDNA: the structural basis for hormone-
RL activated transmembrane signalling."
RN Cell 40:747-758 (1985).
RX 2;
RP SEQUENCE FROM N.A. (ISOFORM SHORT), AND VARIANTS ILE-448 AND GLN-492.
RX MEDLINE=85137889; PubMed=2983222;
RA Ullrich A., Bell J.R., Chen E.Y., Herrera R., Petruzzelli L.M.,
RA Dull T.J., Gray A., Coussens L., Liao Y.-C., Tsibokawa M.,
RA Mason A., Seeburg P.H., Grunfeld C., Rosen O.M., Ramachandran J.;
RT "Human insulin receptor and its relationship to the tyrosine kinase
RL family of oncogenes."
RN Nature 313:756-761 (1985).
RX 3;
RP REVISION TO 899-900.
RA Chen E.Y.;
RL Submitted (JUL-1985) to the EMBL/GenBank/DBJ databases.
RX 4;
RP SEQUENCE FROM N.A. (ISOFORM LONG), AND VARIANTS ILE-448 AND GLN-492.
RX TISSUE=Liver;
RL MEDLINE=91006864; PubMed=2210055;
RA Saino S., Saino M., Bell G.I.;
RT "Human insulin-receptor gene. Partial sequence and amplification of
RL exons by polymerase chain reaction."
RN Diabetes 39:123-128 (1990).
RX 5;
RP SEQUENCE OF TRYPTIC PEPTIDES.
RX MEDLINE=88190050; PubMed=3447155;
RA Fujita-Yamaguchi Y., Hawke D., Shively J.E., Choi S.;
RT "Partial amino acid sequence analyses of human placental insulin
RL receptor."
RN Protein Seq. Data Anal. 1:3-6 (1987).
RX 6;
RP SEQUENCE OF TRYPTIC PEPTIDES.
```

RL MEDLINE=91009374; PubMed=2211730;
RN Xu Q.Y., Paxton R.J., Fujita-Yamaguchi Y.;
RP "Structural analysis of the insulin receptor by microsequence
RT analyses of limited tryptic fragments isolated by sodium dodecyl
RT sulfate-polyacrylamide gel electrophoresis in the absence or presence
RT of dithiothreitol.";
RL J. Biol. Chem. 265:18673-18681(1990).
RN [7]
RP SEQUENCE OF 1-33 FROM N.A.
RX MEDLINE=88058985; PubMed=3680248;
RA Araki E., Shimada P., Uzawa H., Mori M., Ebina Y.;
RT "Characterization of the promoter region of the human insulin
RT receptor gene. Evidence for promoter activity.";
RL J. Biol. Chem. 262:16186-16191(1987).
RN [8]
RP SEQUENCE OF 1-33 FROM N.A.
RX MEDLINE=89380228; PubMed=2777789;
RA Tewari D.S., Cook D.M., Taub R.;
RT "Characterization of the promoter region and 3' end of the human
RT insulin receptor gene.";
RL J. Biol. Chem. 264:16238-16245(1989).
RN [9]
RP SEQUENCE OF 1-33 FROM N.A.
RX MEDLINE=89165872; PubMed=2538124;
RA Seino S., Bell G.I.;
RT "Alternative splicing of human insulin receptor messenger RNA.";
RL Biochem. Biophys. Res. Commun. 159:312-316(1989).
RN [10]
RP SEQUENCE OF 895-1085 FROM N.A.
RX MEDLINE=89252471; PubMed=2566545;
RA Elbein S.C.;
RT "Molecular and clinical characterization of an insertional
RT polymorphism of the insulin-receptor gene.";
RL Diabetes 38:737-743(1989).
RN [11]
RP SEQUENCE OF 1006-1123 FROM N.A.
RX MEDLINE=89298408; PubMed=2544597;
RA Taira M., Taira M., Hashimoto N., Shimada P., Suzuki Y.,
RA Kanatsuka A., Nakamura P., Ebina Y., Tatibana M., Makino H.;
RT "Human diabetes associated with a deletion of the tyrosine kinase
RT domain of the insulin receptor.";
RL Science 245:63-66(1989).
RN [12]
RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITE ASN-541.
RX MEDLINE=93112026; PubMed=1472036;
RA Schaeffer L., Ljungqvist L.;
RT "Identification of a disulfide bridge connecting the alpha-subunits
RT of the extracellular domain of the insulin receptor.";
RL Biochem. Biophys. Res. Commun. 189:650-653(1992).
RN [13]
RP AUTOPHOSPHORYLATION
RX MEDLINE=92337603; PubMed=1321605;
RA Dickens M., Favare J.M.;
RT "Analysis of the order of autophosphorylation of human insulin
RT receptor tyrosines 1158, 1162 and 1163.";
RL Biochem. Biophys. Res. Commun. 186:244-250(1992).
RN [14]
RP MUTAGENESIS OF TYR-999
RX MEDLINE=88311065; PubMed=2842060;
RA White M.F., Livingston J.N., Backer J.M., Lauris V., Dull T.J.,
RA Ullrich A., Kahn C.R.;
RT "Mutation of the insulin receptor at tyrosine 960 inhibits signal
RT transmission but does not affect its tyrosine kinase activity.";
RN [15]
RP MUTAGENESIS OF LYS-1057
RX MEDLINE=87118237; PubMed=3101064;
RA Ebina Y., Araki E., Taira M., Shimada P., Mori M., Craik C.S.,
RA Siddie K., Pierce S.B., Roth R.A., Rutter W.J.;
RT "Replacement of the lysine residue 1030 in the putative ATP-binding
RT region of the insulin receptor abolishes insulin- and
RT antibody-stimulated glucose uptake and receptor kinase activity.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:704-708(1987).
RN [16]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1005-1310.
RX MEDLINE=95089813; PubMed=7997262;
RA Hubbard S.R., Wei L., Ellis L., Hendrickson W.A.;
RT "Crystal structure of the tyrosine kinase domain of the human insulin
RT receptor.";
RL Nature 372:746-754(1994).
RN [17]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1005-1310.
RX MEDLINE=97459943; PubMed=9312016;
RA Hubbard S.R.;
RT "Crystal structure of the activated insulin receptor tyrosine kinase
RT in complex with peptide substrate and ATP analog.";
RL EMBO J. 16:5572-5581(1997).
RN [18]
RP VARIANT INS RESISTANCE SER-762.
RX MEDLINE=88204915; PubMed=3283938;
RA Yoshimasa Y., Seino S., Whitaker J., Kakehi T., Kosaki A., Kuzuya H.,
RA Imura H., Bell G.I., Steiner D.F.;
RT "Insulin-resistant diabetes due to a point mutation that prevents
RT insulin proreceptor processing.";
RL Science 240:784-787(1988).
RN [19]
RP VARIANT LEPRECHAUNISM GLU-487
RX MEDLINE=88204916; PubMed=2834824;
RA Kadowaki T., Bevins C., Cama A., Ojamaa K., Marcus-Samuels B.,
RA Kadowaki H., Beitz L., McKeon C., Taylor S.I.;
RT "Two mutant alleles of the insulin receptor gene in a patient with
RT extreme insulin resistance.";
RL Science 240:787-790(1988).
RN [20]
RP VARIANT LEPRECHAUNISM PRO-260
RX MEDLINE=90060008; PubMed=2495553;
RA Klinkhamer M.P., Groen N.A., van der Zon G.C.M., Lindhout D.,
RA Sandkuyil L.A., Krans H.M.J., Moeller W., Maassen J.A.;
RT "A leucine-to-proline mutation in the insulin receptor in a family
RT with insulin resistance.";
RL EMBO J. 8:2503-2507(1989).
RN [21]
RP VARIANT INS RESISTANCE VAL-1035.
RX MEDLINE=89298409; PubMed=2544998;
RA Odawara M., Kadowaki T., Yamamoto R., Shibasaki Y., Tobu K.,
RA Accili D., Bevins C., Mikami Y., Matsuura N., Akanuma Y., Takaku F.,
RA Taylor S.I., Kasuga M.;
RT "Human diabetes associated with a mutation in the tyrosine kinase
RT domain of the insulin receptor.";
RL Science 245:66-68(1989).
RN [22]
RP VARIANT INS RESISTANCE THR-1161.
RX MEDLINE=90368673; PubMed=2168397;
RA Moller D.E., Yokota A., White M.F., Pazianos A.G., Flier J.S.;
RT "A naturally occurring mutation of insulin receptor alanine 1134
RT impairs tyrosine kinase function and is associated with dominantly
RT inherited insulin resistance.";
RL J. Biol. Chem. 265:14979-14985(1990).
RN [23]
RP CHARACTERIZATION OF VARIANT RABSON-MENDENHALL SYNDROME LYS-42.
RX MEDLINE=91035445; PubMed=2121734;
RA Kadowaki T., Kadowaki H., Accili D., Taylor S.I.;
RT "Substitution of lysine for asparagine at position 15 in the
RT alpha-subunit of the human insulin receptor. A mutation that impairs
RT transport of receptors to the cell surface and decreases the affinity
RT of insulin binding.";
RN [24]


```

FT DISULFID 461 494
FT FT DISULFID 550 550
FT FT CARBOHYD 42 42
FT FT CARBOHYD 51 51
FT FT CARBOHYD 104 104
FT FT CARBOHYD 137 137
FT FT CARBOHYD 241 241
FT FT CARBOHYD 281 281
FT FT CARBOHYD 321 321
FT FT CARBOHYD 363 363
FT FT CARBOHYD 423 423
FT FT CARBOHYD 444 444
FT FT CARBOHYD 540 540
FT FT CARBOHYD 634 634
FT FT CARBOHYD 652 652
FT FT CARBOHYD 699 699
FT FT CARBOHYD 770 770
FT FT CARBOHYD 783 783
FT FT CARBOHYD 921 921
FT FT CARBOHYD 934 934
FT FT CONFLICT 1119 1119
SQ SEQUENCE 1383 AA; 156756 MW; 4B919566902A944A CRC64;

Query Match 11.5%; Score 263.5; DB 1; Length 1383;
Best Local Similarity 24.4%; Pred. No. 1.3e-12;
Matches 108; Conservative 58; Mismatches 151; Indels 125; Gaps 23;

Qy 11 LLLALPEGAAS-----TQVCTGDMKRLPASPETHDMLRLHYQCCVQGNLELYL 65
Db 14 LLMAVAVAGGTAGHLYPEVCPGMDIRNL-----TRL-----HELENCVIRGELQILLM 64

Qy 66 ---PTN-ASLSPQDIQEVGVYLIHNNQVROVPLQRLR-----IVRGTLFEDNYA 113
Db 65 FKTPEDFRDLSFKLLM-ITDYILL-----FRVYGLSLKDLFPNLTVIRGSLF-FNYA 118

Qy 114 LAVLDNGDPLNTPVTGASGGRLRLQLSLTILKGVLIQRNPOLCQDTILWKDI- 172
Db 119 LVIFEMVH-----LKLGLYNLMNITRGSVRIEKNBELCYLATIDMSRIL 163

Qy 173 -FKNNQLALTLDITNRRACHPCSPMKGS-----RCWGSSSDCQ 213
Db 164 DIVSDNTVILNKDNEE-----GCVCPGAKTKTNCPTAVINGQVFERCWTHS--HCQ 215

Qy 214 SLRTVCGAGGCARCKGLPTDCHCEQAAGCTGPKH--SDCLACLHFNHSGICELHCPALV 272
Db 216 KVCPTICKSHGCTAEG-----CCHKECLGNCSEDDPTKVCACRFYLDGQCVETCPPY 271

Qy 273 TYNTD-----TF-----ESMPNPEGRYTFGASCVTACPYNYLSTD-----V 308
Db 272 YHFQDMRCVNFSCQDLHYKCRNRKPGCHQYVIHNNKCIPECPSGGTWSSNLMCTPCL 331

Qy 309 GSCFLVCPLHNOEVTAEADGTQRCCKSKPCA-----RGTHSLLPRAAPVPPFLRM 358
Db 332 GPCPKVQILGEXTIDSVTSAQE--LRGCTVINGSLIINIRGNNL-----AAELEA 382

Qy 359 QPGPAHVLSTL--RPSNDLVS 378
Db 383 NLGLIEISGFLKIRSYALVS 404

```

RESULT 24

INSR_MOUSE

ID INSR_MOUSE

AC P15208;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Insulin receptor precursor (EC 2.7.1.112) (IR).

GN INSR.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBI_TaxId=10090;

```

RN SEQUENCE FROM N.A.
RP MEDLINE=90094325; PubMed=2557333;
RA Flores-Riveros J.R., Sibley E., Kastelic T., Lane M.D.;
RT "Substrate phosphorylation catalyzed by the insulin receptor tyrosine
kinase. Kinetic correlation to autophosphorylation of specific sites
in the beta subunit.";
RL J. Biol. Chem. 264:21557-21572(1989).
RN [2]
RP SEQUENCE OF 1-33 FROM N.A.
RX MEDLINE=90099338; PubMed=2602374;
RA Sibley E., Kastelic T., Kelly T.J., Lane M.D.;
RT "Characterization of the mouse insulin receptor gene promoter.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:9732-9736(1989).
CC -1- FUNCTION: This receptor binds insulin and has a tyrosine-protein
kinase activity.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
CC -1- ENZYME REGULATION: Autophosphorylation activates the kinase
activity.
CC -1- SUBUNIT: Tetramer of 2 alpha and 2 beta chains linked by disulfide
bonds. The alpha chains contribute to the formation of the ligand-
binding domain, while the beta chains carry the kinase domain.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
receptor subfamily.
CC -1- SIMILARITY: Contains 2 fibronectin type III domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
CC -----
DR EMBL; J05149; AAA93318.1; --
DR EMBL; M28869; AAA93319.1; --
DR PIR; A34157; A34157.
DR HSSP; P06213; 1IRK.
DR MGD; MGI:96575; Insr.
DR GO; GO:0009887; P:organogenesis; IMP.
DR InterPro; IPR000494; EGFR_L domain.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow fac recep.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR002011; Recepttyrkinsil.
DR InterPro; IPR001245; Tyr_kinase_AS.
DR InterPro; IPR008286; Tyr_kinase_AS.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; kinase; 1.
DR Pfam; PF01030; Recep_L domain; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00261; FU; 1.
DR SMART; SM00219; TyRK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE TVR; 1.
DR PROSITE; PS00239; RECEPTOR TYR KIN II; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; tyrosine-protein kinase; Receptor; Transmembrane;
KW Glycoprotein; ATP-binding; Phosphorylation; Signal; Repeat.
FT SIGNAL 1 27
FT CHAIN 28 749 INSULIN RECEPTOR, ALPHA-SUBUNIT.
FT PROPEP 750 752 REMOVED IN MATURE FORM.
FT CHAIN 753 1372 INSULIN RECEPTOR, BETA-SUBUNIT.
FT DOMAIN 753 946 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 947 967 POTENTIAL.

```

FT DOMAIN 988 1372 CTOPLASMIC (POTENTIAL).
FT DOMAIN 520 837 FIBRONECTIN TYPE-III 1.
FT DOMAIN 838 938 FIBRONECTIN TYPE-III 2.
FT DOMAIN 1013 1288 PROTEIN KINASE.
FT NP_BIND 1019 1027 ATP (BY SIMILARITY).
FT BINDING 1047 1047 ATP.
FT ACT_SITE 1149 1149 BY SIMILARITY.
FT MOD_RES 1179 1179 PHOSPHORYLATION (AUTO-).
FT ACT_SITE 989 989 IMPORTANT FOR BIOLOGICAL ACTIVITY.
FT DISULFID 219 228 BY SIMILARITY.
FT DISULFID 223 234 BY SIMILARITY.
FT DISULFID 235 243 BY SIMILARITY.
FT DISULFID 239 252 BY SIMILARITY.
FT DISULFID 255 264 BY SIMILARITY.
FT DISULFID 286 293 BY SIMILARITY.
FT DISULFID 301 311 BY SIMILARITY.
FT DISULFID 315 328 BY SIMILARITY.
FT DISULFID 331 335 BY SIMILARITY.
FT DISULFID 462 485 BY SIMILARITY.
FT DISULFID 551 551 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 43 43 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 52 52 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 322 322 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 424 424 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 541 541 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 635 635 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 653 653 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 700 700 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 759 759 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 772 772 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 910 910 N-LINKED (GLUCNAC. .) (POTENTIAL).
FT CARBOHYD 923 923 N-LINKED (GLUCNAC. .) (POTENTIAL).
SQ SEQUENCE 1372 AA; 155639 MM; 1DA2A0BB74618964 CRC64;
Query Match 11.4%; Score 261; DB 1; Length 1372;
Best Local Similarity 24.8%; Pred. No. 28-13;
Matches 107; Conservative 54; Mismatches 147; Indels 124; Gaps 23;
16 LPPGAASQVCTGTGDKMLRPLSPSPETHLMLHLYQGQVQVQGNLELTVL-----PTN-AS 70
29 LYPG-----EVCFGDIRNNL-----TRD-----HELENCVSIEGHLQILLMKTREDPRD 75
71 LSLFQIQEVQVQVLIANQVQVPLQRLR-----IVRGTLFEDNYVALVLDNGDPL 123
76 LSPFKLIM-ITDYLL-----FRVYGLSILKOLFPLNLTVIRGSRFP-FNYALVIFEMVH-- 127
124 NNTTPTVGASPGGLRELQRLSITELKGVLFQRMFQPCYQDTILWKDIFH--KNNQAL 181
128 -----LKEGLGLYMLNITRSGVRIEKNKELCYLATIDMSILDSVEDNYVL 174
182 TLIDNRSGACHPCSPMKCGS-----RCWGESSEDCQSILTRVCAGG 223
175 NKDDNEE-----CGDVCPGTAKGTNCPATVINGQFVERCWTHS--HCQKVCPTICKSH 226
224 CARCKGPLPTDCHRCQCAAGTGPKH--SDCLACLHFNHSGICELHCPALVTYNTD----- 277
227 GCTAEGL-----CCHKECTLGNCEPDPTKCVACRNFYLDGQCVETCPPPYHFQDWRCVN 282
278 -TF-----ESMPNPEGYTFGASCVTACPNYILSTD-----VGSCTLVCLPLR 318
283 PSFCQDLHFKCRNSRKPQCHQVVIHNNKCIPECPSGYTWSSNLMCTPCIGCPKVCQIL 342
319 NQEVTAEDGTQCEKCSKPCA-----RGTHSLIPRAANVPVLEWQFPAHPVLS 368
343 EGEKTIIDSVTSAQE--LRGCTVINGSLIINIRGNNL-----AAELEAMGLIEEISG 393

QY 369 FL--RPSWDLVS 378
DB 394 FLKIRRYALVS 405
RESULT 25
INSTR AEDAE
ID INSR_AEDAE STANDARD; PRT; 1390 AA.
AC Q93105;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Insulin-like receptor precursor (EC 2.7.1.112) (MIR).
OS Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
OX NCBI_TaxID:7159;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UGAL; TISSUE=Ovary;
RX MEDLINE=97254344; PubMed=9099579;
RA Graf R., Neuschwander S., Brown M.R., Ackermann U.;
RT "Insulin-mediated secretion of ecdysteroids from mosquito ovaries and
RT molecular cloning of the insulin receptor homologue from ovaries of
RT bloodfed Aedes aegypti.";
RL Insect Mol. Biol. 6:151-163 (1997).
CC -!- FUNCTION: THIS RECEPTOR PROBABLY BINDS AN INSULIN RELATED PROTEIN
CC AND HAS A TYROSINE-PROTEIN KINASE ACTIVITY (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBUNIT: Tetramer of 2 alpha and 2 beta chains linked by disulfide
CC bonds. The alpha chains contribute to the formation of the ligand-
CC binding domain, while the beta chains carry the kinase domain (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
CC receptor subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

EMBL; U72939; AAB17094.1; --
PIR; T30346; T30346.
HSP; P06213; IIRK
InterPro; IPR000494; EGFR_L domain.
InterPro; IPR003961; FN_III.
InterPro; IPR006211; Furin-like.
InterPro; IPR006212; Furin repeat.
InterPro; IPR009030; Grow fac recep.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR020111; RecepttyrkinasII.
InterPro; IPR001245; Tyr_kinase.
InterPro; IPR008266; Tyr_kinase_AS.
Pfam; PF00041; fn3; 1.
Pfam; PF00757; Furin-like; 1.
Pfam; PF00069; pkinase; 1.
Pfam; PF01030; Recep_L domain; 2.
PRINTS; PR00109; TYRKINASE
ProDom; PD000001; prot_kinase; 1.
SMART; SM00060; FN3; 3.
SMART; SM00261; FU; 1.
SMART; SM00219; Tyrc; 1.
PROSITE; PS00107; PROTEIN KINASE ATP; 1.
PROSITE; PS00109; PROTEIN KINASE TYR; 1.
PROSITE; PS00239; RECEPTOR TYR_KIN II; 1.
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
Glycoprotein; ATP-binding; Phosphorylation; Signal.
KW

```
FT SIGNAL 1 19 POTENTIAL. INSULIN-LIKE RECEPTOR, ALPHA CHAIN
FT CHAIN 20 753 (POTENTIAL).
FT PROPEP 754 757 (POTENTIAL).
FT CHAIN 758 1390 INSULIN-LIKE RECEPTOR, BETA CHAIN
(POTENTIAL).
FT DOMAIN 758 966 EXTRACELLULAR (POTENTIAL).
FT TRANSKEM 967 987 POTENTIAL.
FT DOMAIN 988 1390 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1025 1301 PROTEIN KINASE.
FT NP BIND 1031 1039 ATP (BY SIMILARITY).
FT BINDING 1059 1059 ATP (BY SIMILARITY).
FT ACT SITE 1162 1162 BY SIMILARITY.
FT MOD_RES 1192 1192 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 176 176 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 322 322 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 492 492 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 553 553 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 765 765 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 772 772 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 793 793 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 796 796 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 871 871 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 919 919 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 933 933 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 1390 AA; 156831 MW; 6802AA382EF6442 CRC64;

Query Match 11.3%; Score 258; DB 1; Length 1390;
Best Local Similarity 25.1%; Pred. No. 3.6e-12;
Matches 95; Conservative 53; Mismatches 124; Indels 106; Gaps 17;

Qy 16 LPFGAASCTVCTGDMKRLRSPAPETHLDMLRHLGYQGVQGNLSLTYP-----TN 68
Db 30 VPKGG-----VCGTVQVDR-----NSPAHLDELK-----DCVVVEGFVHLLIDKYIDSSFEN 76
Qy 69 ASLSFLQIOCEVQGVYLIHNRQVPLQRLRIVRGTPQLFEDNVALVLDNGDPLNNTTP 128
Db 77 YSPFLLEITE-----YLL-----FRVNGKSLR-----RLFPN---LAVYP-GDALVGDYA 120
Qy 129 VTGASPGGLRELQRLSRITLKGVLQIQRNPOLCYQDTILWKDI-----FHNQL--- 179
Db 121 MVIYELMHIEIGLISLMDITRGVRIEKNPKLCPANTIDWKAMTVPGTNNYIKDQKXDN 180
Qy 180 -----ALTLIDTNSRACHPCSPMCKGR-----CWGESSECCQSLTRVVCAGGCA 225
Db 181 VCPICPAESTAVMLPNSGKQKCAAPVRGNGKHKTLCW--NANHCQ-----TICPECP 234
Qy 226 RCKGPLPTDCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNP 285
Db 235 KACSKTGVCDDAESCGLGCLNPLNTSSCSVCRHLSIDPAGKRQCVAKCPNPT----- 285
Qy 286 EGRVTFGASCVTA-----CPYNYLSTDVGSCTLVCLPLNQEV 322
Db 286 ---FKYTRCVTRDECTAMKPKPISLSDNPDLDPQPFIPHN-----GSLMECPVDHLLI 336
Qy 323 TAEDGTQRCCKSKPCAR 340
Db 337 TELNKTWRCKNCKSGTCFK 354

RESULT 26
IGIR HUMAN STANDARD; PRT; 1367 AA.
ID IGIR HUMAN AC P08069;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Insulin-like growth factor I receptor precursor (EC 2.7.1.112)
```

```
DE (CD221 antigen).
GN IGFIR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC MEDLINE=87053815; PubMed=2877871;
RX Ullrich A., Gray A., Tam A.W., Yang-Feng T., Taubokawa M.,
RA Collins C., Henzel W., Bon T.L., Kathuria S., Chen E., Jacobs S.,
RA Francke U., Ramachandran J., Fujita-Yamaguchi Y.;
*Insulin-like growth factor I receptor primary structure: comparison
with insulin receptor suggests structural determinants that define
functional specificity.;
EMBO J. 5:2503-2512(1986).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=92268129; PubMed=1316909;
RA Abbot A.M., Bueno R., Pedrini M.T., Murray J.M., Smith R.J.;
*Insulin-like growth factor I receptor gene structure.;
J. Biol. Chem. 267:10759-10763(1992).
[3]
SEQUENCE OF 1-31 FROM N.A.
RX MEDLINE=91282751; PubMed=1711844;
RA Cooke D.W., Bankert L.A., Roberts C.T. Jr., Leroith D.,
RA Casella S.J.;
*Analysis of the human type I insulin-like growth factor receptor
promoter region.;
Biochem. Biophys. Res. Commun. 177:1113-1120(1991).
[4]
SEQUENCE OF 1137-1193 FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=Melanocyte;
RX MEDLINE=94067791; PubMed=8247543;
RA Lee S.-T., Strunk K.M., Spritz R.A.;
*A survey of protein tyrosine kinase mRNAs expressed in normal human
melanocytes.;
Oncogene 8:3403-3410(1993).
CC -!- FUNCTION: THIS RECEPTOR BINDS INSULIN-LIKE GROWTH FACTOR I (IGF I)
WITH A HIGH AFFINITY AND IGF II WITH A LOWER AFFINITY. IT HAS A
TYROSINE-PROTEIN KINASE ACTIVITY.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
CC -!- SUBUNIT: Tetramer of 2 alpha and 2 beta chains linked by disulfide
bonds. The alpha chains contribute to the formation of the ligand-
binding domain, while the beta chain carries the kinase domain.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in a variety of tissues.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
receptor subfamily.
CC -!- SIMILARITY: Contains 2 fibronectin type III domains.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; X04434; CAA28030.1; -.
DR EMBL; M69229; AAB59399.1; -.
DR PIR; A25690; IGHR1.
DR PDB; 1IGR; 27-SEP-99.
DR PDB; 1JQH; 18-DEC-02.
DR PDB; 1K3A; 05-DEC-01.
DR PDB; 1MTN; 14-JAN-03.
DR PDB; HGNC:5465; IGFIR.
DR MIM; 147370; -.
DR GO; GO:0005010; P:insulin-like growth factor receptor activity; TAS.
DR GO; GO:0006916; P:anti-apoptosis; TAS.
DR GO; GO:0008286; P:insulin receptor signaling pathway; TAS.
```


tyrosine phosphate.
 -!- SUBUNIT: Tetramer of 2 alpha and 2 beta chains linked by disulfide bonds. The alpha chains contribute to the formation of the ligand-binding domain, while the beta chain carries the kinase domain.
 -!- SUBCELLULAR LOCATION: Type I membrane protein.
 -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin receptor subfamily.
 -!- SIMILARITY: Contains 2 fibronectin type III domains.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; L29232; AAA41392.1; -;
 EMBL; M27293; AAA41384.1; -;
 HSP; P06213; 1TRK.
 InterPro; IPR000494; EGFR_L domain.
 InterPro; IPR008957; FN III-like.
 InterPro; IPR003961; FN III.
 InterPro; IPR006211; Furin-like.
 InterPro; IPR006212; Furin repeat.
 InterPro; IPR009030; Grow_Fac_recep.
 InterPro; IPR000719; Prot_kinase.
 InterPro; IPR002011; Recepttyr_kinsII.
 InterPro; IPR001245; Tyr_pkinase.
 InterPro; IPR008266; Tyr_pkinase_AS.
 Pfam; PF00041; fn3; 2.
 Pfam; PF00757; Furin-like; 1.
 Pfam; PF00069; pkinase; 1.
 Pfam; PF01030; Recept_L domain; 2.
 PRINTS; PR00109; Recept_L domain.
 ProDom; PD000001; Prot_kinase; 1.
 SMART; SM00060; FN3; 3.
 SMART; SM00261; FU; 1.
 SMART; SM00219; TyrK; 1.
 PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 PROSITE; PS00239; RECEPTOR TYR_KIN_II; 1.
 PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane; Glycoprotein; ATP-binding; Phosphorylation; Repeat; Signal.
 SIGNAL 1 30
 CHAIN 31 741 INSULIN-LIKE GROWTH FACTOR I RECEPTOR,
 ALPHA-CHAIN
 CHAIN 742 1370 INSULIN-LIKE GROWTH FACTOR I RECEPTOR,
 BETA-CHAIN.
 DOMAIN 742 936 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 937 960 POTENTIAL.
 DOMAIN 961 1370 CYTOPLASMIC (POTENTIAL).
 DOMAIN 608 829 FIBRONECTIN TYPE-III 1.
 DOMAIN 830 929 FIBRONECTIN TYPE-III 2.
 DOMAIN 1000 1275 PROTEIN KINASE.
 NP_BIND 1006 1014 ATP (BY SIMILARITY).
 BINDING 1034 1034 ATP (BY SIMILARITY).
 ACT_SITE 1136 1136 BY SIMILARITY.
 DISULFID 215 224 BY SIMILARITY.
 DISULFID 219 230 BY SIMILARITY.
 DISULFID 231 239 BY SIMILARITY.
 DISULFID 235 248 BY SIMILARITY.
 DISULFID 251 260 BY SIMILARITY.
 DISULFID 264 276 BY SIMILARITY.
 DISULFID 282 303 BY SIMILARITY.
 DISULFID 307 321 BY SIMILARITY.
 DISULFID 324 328 BY SIMILARITY.
 CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 535 535 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 641 641 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 748 748 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 757 757 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 765 765 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 901 901 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 914 914 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT MOD_RES 1166 1166 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 985 985 AD -> PY (IN REF. 3).
 FT CONFLICT 1370 AA; 155395 MW; A5946897A1C8145 CRC64;
 SQ SEQUENCE
 Query Match 11.0%; Score 252.5; DB 1; Length 1370;
 Best Local Similarity 23.4%; Pred. No. 9.4e-12;
 Matches 90; Conservative 53; Mismatches 121; Indels 121; Gaps 22;
 QY 9 WGL--LLALLPGGAATOV--TGTDMKRLRPASPETHLDMRLHYOGCVVQGNLELTYL 65
 DB 14 WGLVFLSAALSUWPTSGEICGPODIR-----NDYQQLKRL-ENCIVIEGFHILL- 63
 QY 66 PTNASLFLQDIOE-----VOGYVLIAHNOVRQVPLQ-----RLRIVRGTOLEFED 110
 DB 64 ----ISKAEVRSYRPPKLTIVITEYLL-----FRVAGLESGLDLPENLTIVIRGKLFY- 113
 QY 111 NYALAVLDNGDPLNNTPTVGTASPGGRLQLSLEILKGGVLIORNPQLCYQDITLWK 170
 DB 114 NYALVIFE-----MTN-----LKDGLNLRNITIRGAIRIEKNADLCVLSILDW- 157
 QY 171 DIFHRNNQALTLITDNRSRACHP-----CSPMKKGS-----RCWGESS 209
 DB 158 -----SLILDVAVNNYIVGNKPKKCGDLCPTLEKPKCKTTINNEYRCW--TT 208
 QY 210 EDQSLTRTVGAGGCARCKGPLETDCHEQCACAGCTGP-KHSDCLACLFNHSIGELHC 268
 DB 209 NRCQKQCPSPVC--GKACTE--NNECCHEPCGSGCHTPODNTTCVACRYYYGVCVPAC 264
 QY 269 PALVTY-----NTDTFESMPNPEGRYTFG-----ASCWTACPYNYLSTDVGSCTLVC 315
 DB 265 PP-GTYRFGWRCVDRDFCANIPNAESSDSGVIHDEGCMQSCPSGFRINSTQS----- 318
 QY 316 PLNQEVTAEDGTQRCCKSKPCAR 340
 DB 319 -----MYCIPCEGPCPK 330
 RESULT 28
 IGLR_MOUSE STANDARD; PRT; 1373 AA.
 AC Q60751; O70438; Q62123;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 GN Insulin-like growth factor I receptor precursor (BC 2.7.1.112).
 OS IGLR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Navarro M., Garandel V., Barenton B., Bernardi H.;
 RA "Cloning of cDNA for the mouse insulin-like growth factor I
 RT receptor.";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RP SEQUENCE OF 1-329 FROM N.A.
 RC STEIN-CD-1; TISSUE=Kidney;
 RA Jun W., Liu Z., Alvares K., Kumar A., Wallner E.I., Kanwar Y.S.;
 RA Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]

```
RP SEQUENCE OF 1134-1203 FROM N.A.
```

```
RA MIDDLELINE=90152381; PubMed=2462828;
```

```
RK Wilks A.F., Kurban R.R., Hovens C.M., Ralph S.J.;
```

```
RT "The application of the polymerase chain reaction to cloning members
```

```
RL of the protein tyrosine kinase family.";
```

```
RL Gene 85:67-74(1983).
```

```
CC -!- FUNCTION: THIS RECEPTOR BINDS INSULIN-LIKE GROWTH FACTOR I (IGF I)
```

```
CC WITH A HIGH AFFINITY AND IGF II WITH A LOWER AFFINITY. IT HAS A
```

```
CC TYROSINE-PROTEIN KINASE ACTIVITY.
```

```
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
```

```
CC tyrosine phosphate.
```

```
CC -!- SUBUNIT: Tetramer of 2 alpha and 2 beta chains linked by disulfide
```

```
CC bonds. The alpha chains contribute to the formation of the ligand-
```

```
CC binding domain while the beta chain carries the kinase domain.
```

```
CC -!- Subcellular location: Type I membrane protein.
```

```
CC -!- Similarity: Belongs to the tyr family of protein kinases. Insulin
```

```
CC receptor subfamily.
```

```
CC -!- Similarity: Contains 2 fibronectin type III domains.
```

```
-----
```

```
This SWISS-Prot entry is copyright. It is produced through a collaboration
```

```
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
```

```
the European Bioinformatics Institute. There are no restrictions on its
```

```
use by non-profit institutions as long as its content is in no way
```

```
modified and this statement is not removed. Usage by and for commercial
```

```
entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licensesh@isb-sib.ch).
```

```
-----
```

```
EMBL; AF056187; AAC12782.1; --
```

```
EMBL; U00182; AAC52123.1; --
```

```
EMBL; M33422; AAA40013.1; --
```

```
FIR; A48805; A48805.
```

```
HSSP; P06213; 1IHK.
```

```
MGI; MG1:96433; Igf1r.
```

```
GO; GO:0009887; Proteinogenesis; IMP.
```

```
InterPro; IPRO00494; EGFR_L domain.
```

```
InterPro; IPRO08957; FN_III-like.
```

```
InterPro; IPRO03961; FN_III.
```

```
InterPro; IPRO06211; Furin-like.
```

```
InterPro; IPRO09030; Furin repeat.
```

```
InterPro; IPRO009030; Grow fac recep.
```

```
InterPro; IPRO00719; Prot_kinase.
```

```
InterPro; IPRO02011; RecepttyrkinsII.
```

```
InterPro; IPRO01245; Tyr_pkinase.
```

```
InterPro; IPRO08266; Tyr_pkinase_AS.
```

```
Pfam; PF00941; fn3; 2.
```

```
Pfam; PF00757; Furin-like; 1.
```

```
Pfam; PF00669; pkinase; 1.
```

```
Pfam; PF01030; Recep_L domain; 2.
```

```
PRINTS; PR00109; TYRKINASE.
```

```
ProDom; PD0000001; Prot_kinase; 1.
```

```
SMART; SM00060; FN3; 3.
```

```
SMART; SM00261; FU; 1.
```

```
SMART; SM00219; TyrcK; 1.
```

```
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
```

```
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
```

```
PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
```

```
PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
```

```
KW Transfrase; Tyrocin-protein kinase; Receptor; Transmembrane;
```

```
KW Glycoprotein; Atfp-binding; Phosphorylation; Repeat; Signal.
```

```
SIGNAL 1 30
```

```
CHAIN 31 741
```

```
INSULIN-LIKE GROWTH FACTOR I RECEPTOR,
```

```
ALPHA-CHAIN.
```

```
BETA-CHAIN.
```

```
EXTRACELLULAR (POTENTIAL).
```

```
POTENTIAL.
```

```
CYTOPLASMIC (BY SIMILARITY).
```

```
FIBONECTIN TYPE-II 1.
```

```
FIBONECTIN TYPE-III 2.PROTEIN KINASE
```

```
DOMAIN 1000 1276
```

```
NB_BIND 1006 1014
```

```
BINDING 1034 1034
```

```
ACT_SITE 1137 1137
```

```
ATP (BY SIMILARITY).
```

FT	DISULFID	215	224	BY SIMILARITY.
FT	DISULFID	219	230	BY SIMILARITY.
FT	DISULFID	231	239	BY SIMILARITY.
FT	DISULFID	235	248	BY SIMILARITY.
FT	DISULFID	251	260	BY SIMILARITY.
FT	DISULFID	264	276	BY SIMILARITY.
FT	DISULFID	282	303	BY SIMILARITY.
FT	DISULFID	307	321	BY SIMILARITY.
FT	DISULFID	324	328	BY SIMILARITY.
FT	CARBOHYD	51	51	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	102	102	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	135	135	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	245	245	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	314	314	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	418	418	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	439	439	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	535	535	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	608	608	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	623	623	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	641	641	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	748	748	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	757	757	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	785	785	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	901	901	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	914	914	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	MOD. RES	1167	1167	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	CONFLICT	58	59	FL -> LV (IN REF. 2).
FT	CONFLICT	260	260	C -> S (IN REF. 2).
FT	CONFLICT	301	301	D -> G (IN REF. 2).
FT	CONFLICT	306	306	E -> V (IN REF. 2).
FT	CONFLICT	324	324	C -> S (IN REF. 2).
FT	CONFLICT	1134	1134	V -> I (IN REF. 3).
FT	CONFLICT	1145	1145	V -> I (IN REF. 3).
FT	CONFLICT	1202	1202	V -> I (IN REF. 3).
FT	SEQUENCE	1373 AA;	155787 MW; 5BE3B72BF101E379 CRC64;	

Query Match	11.0%;	Score	250.5;	DB 1;	Length	1373;	
Best Local Similarity	23.4%;	Pred. No.	1.3e-11;				
Matches	90;	Conservative	53;	Mismatches	121;	Gaps	22;

QY	9	WGL--LLALLPPCAASTVC--TGTDMKRLPASPTHLDMLRHLYOGCQVVOGNLELTYL	65
DB	14	WGLVFLSAALSFWPTSGEIGCGFIDIR-----NDYQQLKEL-ENCIVIEGFLHLL--	63
QY	66	PTNASLSFLQDTQE-----VQGVLIANHNOVRQVPLQ-----RLRIVRGTQLPED	110
DB	64	-----ISRAEDYSRYRFPKLTIVITEYLL-----FRVAGLESGLDLPNPTVIRGMKLFY--	113
QY	111	NYALAVLDNGDPLNNTPTVTGASPGCLRQLRLSLTEILKGVLIORNPOLCVQDITLWK	170
DB	114	NYALVIFE-----MTN-----LKDIGHYLNRTIRGAIRIEKXADLCYLSIDW--	157
QY	171	DIFHKNNQALTLIDNRSRACHP---CSPMCKGS-----RCWGESS	209
DB	158	-----SLILDVSNNYIVGNKPKPEQDLCPTLEEXPMCKEKTINNEYRCW--FT	208
QY	210	EDCQSLRTVCAAGCARCKGPLPTCCHEQCAAGTGP--KESDCLACLHFNHSGICELHC	268
DB	209	NFCQRMCPSPV---GKRACTE---NNBCCHPECIGSCHTTPDNTTCVACRIHYHYKGVCPAC	264
QY	269	PALVTY-----NTDTFESPNPEGRYTFG-----ASCVTACPNYILSTDVGSCTLVC	315
DB	265	PP-GTYRPEGRCVDRDPFCANIPNAESSDSDGFVIHDDSCMCQPCSGFIRNSTQS-----	318
QY	316	PLHNQEVTAEDGTQCEKCKSFCA	340
DB	319	-----MYCIPCEGFCPK	330

RESULT 29
PKCS_BRACL
ID PKCS BRACL
AC Q9NJ15; Q9NJ14; Q9NJ16;
PR7: 1696 AA.

RESULT 29
PCK5_BRACK
ID_PCK5
AC Q9N5

16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Protein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)
PC6
Branchiostoma californiensis (California lancelet) (Amphioxus).
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
NCBI_TaxID=7738;
[1]
SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
MEDLINE=20175281; PubMed=10708869;
Oliva A.A. Jr., Chan S.J., Steiner D.F.;
"Evolution of the prohormone convertases: identification of a
homologue of PC6 in the protochordate amphioxus.";
Biochem. Biophys. Acta 1477:338-348(2000).
-!- FUNCTION: Likely to represent a widespread endoprotease activity
within the constitutive and regulated secretory pathway. Capable
of cleavage at the RX(K/R)R consensus motif (By similarity).
-!- CATALYTIC ACTIVITY: Release of mature proteins from their
propeptides by cleavage of Arg-Xaa-Yaa-Arg-|-Zaa bonds, where Xaa
can be any amino acid and Yaa is Arg or Lys.
-!- SUBCELLULAR LOCATION: ISOFORM A AND ISOFORM C ARE SECRETED.
ISOFORM B IS A TYPE I MEMBRANE PROTEIN.
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=3;
Name=B;
Name=C;
IsoId=Q9NUJ15-1; Sequence=VSP_005444, VSP_005445;
Name=A;
IsoId=Q9NUJ15-2; Sequence=VSP_005444, VSP_005445;
Name=C;
IsoId=Q9NUJ15-3; Sequence=VSP_005442, VSP_005443;
-!- DOMAIN: The propeptide domain acts as an intramolecular chaperone
assisting the folding of the zymogen within the endoplasmic
reticulum.
-!- SIMILARITY: Belongs to peptidase family S8.
-!- SIMILARITY: Contains 1 homo B/P domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF184615; AAF26300.1; -
EMBL; AF184616; AAF26301.1; -
EMBL; AF184617; AAF26302.1; -
HSSP; Q9405; IMPT.
MEROPS; S08.02B; -
InterPro; IPR006212; Furin repeat.
InterPro; IPR009030; Grow fac recep.
InterPro; IPR002029; Peptidase S8.
InterPro; IPR002384; Peptidase_S8.
InterPro; IPR009020; Protease_Inhib.
Pfam; PF01483; P_protease; 1.
Pfam; PF00082; Peptidase_S8; 1.
PRINTS; PR00723; SUBTILISIN.
ProDom; PD000717; P domain; 1.
SMART; SM00261; FU_17.
DR PROSITE; PS00136; SUBTILASE ASP; FALSE_NEG.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
Cleavage on pair of basic residues; Repeat; Alternative splicing;
Transmembrane.
FT SIGNAL 1 25 POTENTIAL.
FT PROPEP 26 110 POTENTIAL.
FT CHAIN 111 1696 PROTEIN CONVERTASE SUBTILISIN/KEXIN
TYPE 5.
FT DOMAIN 111 1618 EXTRACELLULAR (POTENTIAL).

TRANSMEM 1619 1639 POTENTIAL.
DOMAIN 1640 1696 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 111 488 CATALYTIC.
FT DOMAIN 496 637 HOMO B.
FT DOMAIN 664 1649
SITE 110 111
FT ACT_SITE 192 192 CYS-RICH MOTIF (CRM) REGION.
FT ACT_SITE 233 233 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT ACT_SITE 407 407 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 407 407 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 246 246 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 529 529 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 885 885 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARSPLIC 1259 1323 DDTILDGECITSCGGYMDRREKCKCHPTCKCSDBY
PREVSVLAELALHGLHSLYSLTDVPPQSNPPDTVLGADRRL
TTATSAAGRCA (in isoform C).
FT VARSPLIC 1324 1696 Missing (in isoform C).
FT VARSPLIC 1288 1343 /FTIGVSP_005443.
FT VARSPLIC 1344 1696 /FTIGVSP_005445.
FT VARSPLIC 1696 AA; 188410 MM; 281CBEL1784257CBD CRC64;
SQ SEQUENCE 7.4%; Score 169; DB 1; Length 1696;
Query Match Best Local Similarity 25.4%; Pred. No. 3.8e-05;
Matches 64; Conservative 28; Mismatches 106; Indels 54; Gaps 14;
QY 147 EILKGVLIQNPQL---CYQDILWKD-IFKNNQALALITNRSRACHPCSPCKGKS 202
Db 1018 ECAEGCHSCBEGPDICDSCDEYLTEDTCVRRINCPSFTYPD-DQDRECRPCHNCEA- 1075
QY 203 RCWGESSEDCQSL-----TRTVACGACAR--CKGPLETDC--CHEQAAAGCTGPKHSDC 252
Db 1076 -CDGNNGNCSCKEGFYKTPDCSTGCPNRYKDDTNKCKPCDSSCFT-CSGPASFFHC 1133
QY 253 LACL--HFNRSG:CELECPALVYNTDT---FSGMNPBGRYTFGASCTVACPYNYLSTD 307
Db 1134 LSCADGDFLHSSCRSTCPAGFTGNAESHECVSSCEQDQYYSSETGRCDCPCYNCRACD 1193
QY 308 -VGSCTLVCPH-----NQVTAEDGT-----QRCEKCKSP----- 337
Db 1194 NQGDCAECATYIVDGRCPETCDGEYQDRDRDTAELSCRCHQSKCTCGSPDTC 1253
QY 338 -CARGTHSLLR 348
Db 1254 DSCKGDDTILDR 1265
PCK5 MOUSE STANDARD; PRT; 1877 AA.
AC Q04592; Q62040;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Proprotein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)
DE {Proprotein convertase subtilisin/kexin-like protease PCS} (PC6)
DE {Subtilisin-like proprotein convertase 6} (SPC6).
CN PCSK5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OK NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 330-1877 FROM N.A. (ISOFORM PC5B).
RC STRAIN=ICR; TISSUE=Intestine;
RX MEDLINE=93327934; PubMed=8335106;
RA Nakagawa T., Murakami K., Nakayama K.;


```
FT CARBOHYD 951 951 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1016 1016 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1220 1220 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1317 1317 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1523 1523 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1711 1711 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1733 1733 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 978 915 GEYDDQGHGCTCEACXKMGPTQEDCISCFVTRVLD -->
ATEESWAGGFCMLVKNNLCQKVKVLOQLCKCTCTFQG
(in isoform PC5A).
/FTid=VSP 005438.
FT VARSPLIC 916 1877 Missing (in isoform PC5A).
/FTid=VSP 005439.
SQ SEQUENCE 1877 AA; 209287 MW; EC850E2DF0EAIC3 CRC64;

Query Match 6.3%; Score 144; DB 1; Length 1877;
Best Local Similarity 25.2%; Pred. No. 0.0037;
Matches 40; Conservative 14; Mismatches 55; Indels 50; Gaps 9;

Qy 192 CHPCSPKCKGRCKGESSEDCQS--LRTTV---CAGGCARCKGFLPTDC--CHEQCAAG 243
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
887 CQTEASC--AKCWGPTQEDCISCFVTRVLDGRCVWNCPSWKFPEKQCHPCHYC-QG 943
Qy 244 CTGPKHSDCLACHFNHSGICELCPALVTVNTDTFESMPNPEGRYTFGASCVTACPNY 303
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
944 CQSGSPNCTSCRADKHG-----QERFLYHGECLENCVPVGH 979
Qy 304 LSTDVGSCTVACPLHNQEVTAEDGTQCEKCKSP--CAR 340
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
980 YPAGHTC-LPCP-----DNCELCYNPHICSR 1005
```

Search completed: July 4, 2004, 04:19:22
Job time : 24.1928 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: July 4, 2004, 04:15:09 ; Search time 74.8815 Seconds
(without alignments)
1763.484 Million cell updates/sec

Title: US-09-506-079H-12
Perfect score: 2287
Sequence: 1 MELAALCRGMLLALLPPGA.....VGRGPPDPAHVAVNLRSYGS 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 90 summaries

- Database :
- 1: SPTEMBL_25:*
 - 2: sp_archaea:*
 - 3: sp_bacteria:*
 - 4: sp_fungi:*
 - 5: sp_human:*
 - 6: sp_invertebrate:*
 - 7: sp_mammal:*
 - 8: sp_mhc:*
 - 9: sp_organelle:*
 - 10: sp_phase:*
 - 11: sp_plant:*
 - 12: sp_rodent:*
 - 13: sp_virus:*
 - 14: sp_unclassified:*
 - 15: sp_rvrius:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2275	99.5	419	Q9UK79	Q9uk79 homo sapien
2	1716.5	75.1	1259	O18735	O18735 canis famil
3	1608.5	70.3	1259	Q8K3F9	Q8k3f9 rattus norv
4	1597.5	69.9	711	Q80Y89	Q80y89 mus musculu
5	942.5	41.2	431	Q7SY19	Q7sy19 brachydanio
6	794	34.7	527	Q30836	Q30836 gallus gall
7	789	34.5	643	Q3ERV6	Q3erv6 mus musculu
8	789	34.5	655	Q3WVF5	Q3wvf5 mus musculu
9	789	34.5	1210	Q3EP98	Q3ep98 mus musculu
10	788	34.5	1209	Q8MIL8	Q8mil8 sus scrofa
11	781.5	34.2	1209	Q9QX70	Q9qx70 rattus norv
12	773	33.8	478	Q1QE5Q	Q1qe5q rattus norv
13	757.5	33.1	1191	Q7SZF7	Q7szf7 brachydanio
14	723	32.1	331	Q9BUD7	Q9bud7 homo sapien
15	723	31.6	149	Q9BG66	Q9bg66 oryctolagus
16	713.5	31.2	1305	Q8AW81	Q8aw81 brachydanio

17	693	30.3	1165	13	Q9YH40	Q9yh40 xiphophorus
18	662.5	29.0	1328	13	P79754	P79754 fugu rubrip
19	647	28.3	599	13	Q9PSH2	Q9psh2 gallus gall
20	601	26.3	1377	5	Q8MLW0	Q8mlw0 drosophila
21	576.5	25.2	1322	5	Q8GN22	Q8gn22 drosophila
22	571.5	25.0	1433	5	Q9BIH9	Q9bih9 anopheles g
23	463.5	20.3	1137	13	Q9W6F6	Q9w6f6 gallus gall
24	404.5	17.7	150	6	Q9BG64	Q9bg64 oryctolagus
25	402.5	17.6	1368	5	Q23821	Q23821 caenorhabdi
26	395	17.3	151	6	Q9BG65	Q9bg65 oryctolagus
27	363.5	15.9	366	5	Q26569	Q26569 schistosoma
28	363.5	15.9	1717	5	Q26566	Q26566 schistosoma
29	353.5	15.5	1193	5	Q9Y1X8	Q9y1x8 ephydatia f
30	331	14.5	334	5	Q26567	Q26567 schistosoma
31	331	14.5	342	5	Q26568	Q26568 schistosoma
32	324	14.2	1564	5	Q86MD7	Q86md7 echinococcu
33	300.5	13.1	1418	13	Q8UW83	Q8uw83 paralichthy
34	300	13.1	1472	5	Q9U5A8	Q9u5a8 bombyx mori
35	298.5	13.1	1671	5	Q9NJV5	Q9njv5 biophalar
36	298	13.0	1358	13	Q73798	Q73798 xenopus lae
37	298	13.0	1418	13	Q93457	Q93457 scopthalmu
38	294	12.9	2144	5	Q9VD94	Q9vd94 drosophila
39	289	12.6	524	11	Q8C4F9	Q8c4f9 mus musculu
40	284	12.4	1362	13	Q9PV24	Q9pv24 xenopus lae
41	280	12.2	1749	5	Q8TOW6	Q8tow6 echinococcu
42	278	12.2	1412	13	Q8UW84	Q8uw84 paralichthy
43	275.5	12.0	1369	13	Q8UW86	Q8uw86 paralichthy
44	273.5	12.0	89	11	Q88459	Q88459 mus musculu
45	272	11.9	469	11	Q83721	Q83721 rattus norv
46	270	11.8	410	11	Q83720	Q83720 rattus norv
47	263.5	11.5	743	4	Q86WY9	Q86wy9 homo sapien
48	256.5	11.2	1368	13	Q8UW85	Q8uw85 paralichthy
49	252.5	11.0	1371	11	Q9QVW4	Q9qvW4 rattus sp.
50	252	11.0	1499	5	Q86L28	Q86l28 schistosoma
51	232	10.1	1245	13	Q9YGH8	Q9ygh8 scopthalmu
52	231.5	10.1	881	11	Q8CDB7	Q8cdB7 mus musculu
53	230	10.1	946	5	Q9VJ04	Q9vj04 drosophila
54	210	9.2	868	5	Q9VFE2	Q9vfe2 drosophila
55	203	8.9	1843	5	Q968Y9	Q968y9 caenorhabdi
56	203	8.9	1846	5	O16131	O16131 caenorhabdi
57	191	8.4	82	6	Q9N0K4	Q9n0k4 sus scrofa
58	189	8.3	131	5	Q9BH16	Q9bh16 anopheles g
59	181.5	7.9	191	11	Q88458	Q88458 mus musculu
60	171.5	7.5	116	6	Q9BG67	Q9bg67 oryctolagus
61	160.5	7.0	1299	5	Q26489	Q26489 spodoptera
62	156.5	6.8	441	4	Q96JW7	Q96jw7 homo sapien
63	156.5	6.8	651	4	Q86U24	Q86u24 homo sapien
64	156.5	6.8	4007	4	Q86XX4	Q86xx4 homo sapien
65	150	6.6	2327	13	Q3IBG7	Q3ibg7 xenopus lae
66	146.5	6.4	1997	10	Q8LRM7	Q8lrm7 chlamydomon
67	146	6.4	62	6	Q8HY60	Q8hy60 sus scrofa
68	143.5	6.3	4010	11	Q80T14	Q80t14 mus musculu
69	142.5	6.2	39	13	Q9PVG7	Q9pvG7 xiphophorus
70	140.5	6.1	626	4	Q8ND91	Q8nd91 homo sapien
71	140.5	6.1	969	4	Q96KG6	Q96kg6 mus sapien
72	139.5	6.1	747	11	Q8VHP4	Q8vhf4 mus musculu
73	139.5	6.1	1004	11	Q8CGA7	Q8cga7 mus musculu
74	139.5	6.1	1034	11	Q8VHL7	Q8vhl7 mus musculu
75	139.5	6.1	1034	11	Q8VIX5	Q8vik5 mus musculu
76	139	6.1	149	11	O60494	O60494 cavia sp. e
77	137.5	6.0	422	4	Q96HD1	Q96hd1 homo sapien
78	136.5	6.0	1045	5	Q8T3A6	Q8t3a6 caenorhabdi
79	136.5	6.0	1070	5	Q8T3A7	Q8t3a7 caenorhabdi
80	136.5	6.0	1111	5	Q9XWD6	Q9xwd6 caenorhabdi
81	134.5	5.9	815	5	Q16970	Q16970 aplysia cal
82	134.5	5.9	915	11	Q91VK0	Q91vk0 mus musculu
83	133.5	5.8	420	11	Q91XD7	Q91xd7 mus musculu
84	133.5	5.8	824	5	Q16962	Q16962 aplysia cal
85	133.5	5.8	921	11	Q80T91	Q80t91 mus musculu
86	132.5	5.8	947	11	Q8BKK7	Q8bkk7 mus musculu
87	132	5.8	176	11	Q923V5	Q923v5 rattus norv
88	132	5.8	1574	11	O88281	O88281 rattus norv
89	129.5	5.7	417	4	Q9Y409	Q9y409 homo sapien

90 129.5 5.7 913 4 Q96EP4 Q96ep4 homo sapien

ALIGNMENTS

RESULT 1

Q9UK79 PRELIMINARY; PRT; 419 AA.
AC Q9UK79;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Herstatin.
GN HER-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99415951; PubMed=10485918;
RA Doherty J.K., Bond C., Jardim A., Adelman J.P., Clinton G.M.;
OC "The HER-2/neu receptor tyrosine kinase gene encodes a secreted
RT autoinhibitor."
RL Proc. Natl. Acad. Sci. U.S.A. 96:10869-10874 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Doherty J.K., Clinton G.M., Adelman J.P., Evans A.J., Henner W.D.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF177761; AAD56009.2; -
DR GO; GO:0005576; C:extracellular; IDA.
DR GO; GO:0004715; F:non-membrane spanning protein tyrosine kina. . . ; IDA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . . ; IDA.
DR InterPro; IPR000494; EGFR_L domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR009030; Grow_fac_recep.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L domain; 1.
DR SMART; SM00261; FU; 1.
SQ SEQUENCE 419 AA; 45472 MW; FECL1BE347E2D030C CRC64;

Query Match 99.5%; Score 2275; DB 4; Length 419;
Best Local Similarity 99.5%; Pred. No. 2e-201;
Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MELAALCRWGLLLALLPPGAASSTVCTGDMKRLPASPETHLMLRLHYQGCVVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASSTVCTGDMKRLPASPETHLMLRLHYQGCVVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVVLIHNVQVPLQRLIRIVRGTLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVVLIHNVQVPLQRLIRIVRGTLFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTCASPGGLRELQRLSLTEILKGGVLIQNPOLCYQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTTPTVTCASPGGLRELQRLSLTEILKGGVLIQNPOLCYQDTILWKDIFHKNQOLA 180
QY 181 LTLIDTNRSRACHPCSPCKSGRCWGESSEDCQSLTRTVACGACRCKGPLEPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPCKSGRCWGESSEDCQSLTRTVACGACRCKGPLEPTDCCHQC 240
QY 241 AAGCTGPGHSDCLACHFNHSGICELHCPALVTNTDTFESMPNPEGYTFGASCVTACP 300
DB 241 AAGCTGPGHSDCLACHFNHSGICELHCPALVTNTDTFESMPNPEGYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTVLCPLNQEVTADGTQRCCKSPKARGTHSILLPRPAAVPLRMQP 360
DB 301 YNYLSTDVGSCTVLCPLNQEVTADGTQRCCKSPKARGTHSILLPRPAAVPLRMQP 360
QY 361 GPAPVLSFLRPSMDLVSAFVSLPLAPLSPTSPVPSVSGRGDPAHVAVDLSRYEG 419

Db 361 GPAPVLSFLRPSMDLVSAFVSLPLAPLSPTSPVPSVSGRGDPAHVAVDLSRYEG 419

RESULT 2

O18735 PRELIMINARY; PRT; 1259 AA.
AC O18735;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE ErbB-2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Yokota H.;
RL "CDNA cloning of erbB-2 from canine mammary gland."
RT Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB008451; BAA23127.1; -
DR HSP; P11362; IFGK.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . . ; IEA.
DR InterPro; IPR002048; EF-hand
DR InterPro; IPR000494; EGFR_L domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow_fac_recep.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR InterPro; IPR004019; YNP_motif.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; kinase; 1.
DR Pfam; PF01030; Recep_L domain; 2.
DR Pfam; PF02757; YLP; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 3.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00031; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KN ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 1259 AA; 137989 MW; E37364D49C4ACD46 CRC64;

Query Match 75.1%; Score 1716.5; DB 6; Length 1259;
Best Local Similarity 82.2%; Pred. No. 3e-149;
Matches 324; Conservative 14; Mismatches 49; Indels 7; Gaps 2;
QY 1 MELAALCRWGLLLALLPPGAASSTVCTGDMKRLPASPETHLMLRLHYQGCVVQGNL 60
DB 1 MELAALCRWGLLLALLPPGAASSTVCTGDMKRLPASPETHLMLRLHYQGCVVQGNL 60
QY 61 ELTYLPTNASLSFLQDIQEVQGVVLIHNVQVPLQRLIRIVRGTLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIQEVQGVVLIHNVQVPLQRLIRIVRGTLFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTCASPGGLRELQRLSLTEILKGGVLIQNPOLCYQDTILWKDIFHKNQOLA 180
DB 121 DPLNNTTPTVTCASPGGLRELQRLSLTEILKGGVLIQNPOLCYQDTILWKDIFHKNQOLA 180
QY 181 LTLIDTNRSRACHPCSPCKSGRCWGESSEDCQSLTRTVACGACRCKGPLEPTDCCHQC 240
DB 181 LTLIDTNRSRACHPCSPCKSGRCWGESSEDCQSLTRTVACGACRCKGPLEPTDCCHQC 240

QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300
 Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTSCP 300
 QY 301 YNYLSTDVSGCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 355
 Db 301 YNYLSTDVSGCTLVCPHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 355
 QY 356 LRMQPG--PAHVPISFLRPSMDLVSAFYSLPLAP 387
 Db 361 IQEFAGCKIFGLSLAFPLPSFDGDPASNTAPLOP 394

RESULT 3

Q8K3F9 ID Q8K3F9 PRELIMINARY; PRT; 1259 AA.
 AC Q8K3F9;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Neu protooncogene.
 DE Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BDIX;
 RA Watson P.A., Kim K., Chen K.-S., Gould M.N.;
 RT "Androgen-Dependent Mammary Carcinogenesis in Rats Transgenic for the
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY116182; AAM50093.1; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0008468; P:protein amino acid phosphorylation; IEA.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR000494; EGFR L domain.
 DR InterPro: IPR006211; Furin-like.
 DR InterPro: IPR009030; Growth fac. recep.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002290; Ser thr. pkinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR InterPro: IPR008266; Tyr_kinase_AS.
 DR InterPro: IPR004019; YLP motif.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF00069; Pkinase; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR Pfam: PF02757; YLP; 2.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TK; 1.
 DR SMART; SM00219; TYRK; 1.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 1259 AA; 139102 MW; 8724BD5CC33AE953 CRC64;

Query Match 70.3%; Score 1608.5; DB 11; Length 1259;
 Best Local Similarity 85.0%; Pred. No. 2.8e-139;
 Matches 294; Conservative 15; Mismatches 36; Indels 1; Gaps 1;
 QY 1 MELALCRWGLLLALPPGAASVTCVGTGDMKLRPLASPETHLDMLRLHYQCCVVGQNL 60
 Db 4 MELAAWCRWGLLLALPPGIAGTVCVGTGDMKLRPLASPETHLDMLRLHYQCCVVGQNL 63

RESULT 4

Q80Y89 ID Q80Y89 PRELIMINARY; PRT; 711 AA.
 AC Q80Y89;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE V-erb-b2 erythroblastic leukemia viral oncogene homolog 2,
 DE neuro/glioblastoma derived oncogene homolog (Hypothetical
 DE protein).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6; TISSUE=Brain;
 EX MEDLINE=23389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Heiton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC046811; AAH46811.1; -
 DR EMBL; BC053078; AAH53078.1; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0005006; P:epidermal growth factor receptor activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; IEA.
 DR InterPro; IPR000494; EGFR_L domain.
 DR InterPro; IPR006213; Furin-like.
 DR InterPro; IPR006212; Furin-repeat.
 DR InterPro; IPR009030; Grow_fac_recep.
 DR Pfam; PF00757; Furin-like; 1.
 DR SMART; SM00261; FU; 4.
 XW Hypothetical protein.
 SQ SEQUENCE 711 AA; 78707 MW; 682B188EB0E71318 CRC64;

Query Match 69.9%; Score 1597.5; DB 11; Length 711;
 Best Local Similarity 85.6%; Pred. No. 1.4e-138;
 Matches 291; Conservative 15; Mismatches 33; Indels 1; Gaps 1;
 QY 1 MELALCAGWGLLALLPGCAASTOVCTGTDMLRLPASPETHLDMRLHYOCQVVOGVL 60
 DB 1 MELAWCAGFLLALLPGCACTOVCTGTDMLRLPASPETHLDMRLHYOCQVVOGVL 60
 QY 61 ELTYLPNASIFLQDIOEVQGYVLIANQVQVLPQRLRIVRGTLFEDNYVALVLDNG 120
 DB 61 ELTYLPNASIFLQDIOEVQGYVLIANQVQVLPQRLRIVRGTLFEDNYVALVLDNR 120
 QY 121 DPLANN-TTPVTGASPGGLRELQRLSLTEILKGVLIQVLPQCYODTILWKDIFHKNOL 179
 DB 121 DPLENVTAAAPTPEGGLRELQRLSLTEILKGVLIQVLPQCYODVILWKDVLKNNQL 180
 QY 180 ALTLIDNRSRACHPCSPKGRGWGESSEDCQSLTRTVCAAGGCARCKGFLPTDCCHQ 239
 DB 181 APVDMIDNRSRACHPCPCAKDNHGWGESPEDCQILGTICTSGCARCKGFLPTDCCHQ 240
 QY 240 CAAGCTGPKSGDCLACHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTAC 299
 DB 241 CAAGCTGPKSGDCLACHFNHSGICELHCPALVTYNTDTFESMLNPEGRYTFGASCVTTC 300
 QY 300 PNYLSTVGSCTLVCPHNEQVTAEDGTQRCCKSKFCA 339
 DB 301 PNYLSTVGSCTLVCPHNEQVTAEDGTQRCCKSKFCA 340

RESULT 5
 Q7SY19 PRELIMINARY; PRT; 431 AA.
 AC Q7SY19;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]_TaxID=7955;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Body;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman N., Madan A., Rodrigues S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT *Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Body;
 RA Strausberg R.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC055160; AH55160.1; --
 KW Hypothetical protein.
 SQ SEQUENCE 431 AA; 47108 MW; 8326174E46AECBIA CRC64;
 Query Match 41.2%; Score 942.5; DB 13; Length 431;
 Best Local Similarity 54.3%; Pred. No. 2.2e-78;
 Matches 185; Conservative 43; Mismatches 98; Indels 15; Gaps 5;
 QY 9 WGLLLALLPGCAASTOVCTGTDMLRLPASPETHLDMRLHYOCQVVOGKLELYLPTN 68
 DB 11 WVLLLLIGITAAATGRCVCLGTDMLKLPSSLENHYEMRLLYTCQVVGKLEITHLQGN 70
 QY 69 ASLSFLQDIOEVQGYVLIANQVQVLPQRLRIVRGTLFEDNYVALVLDNGDPLNNTTP 128
 DB 71 PDLISFLQDIOEVQGYVLIANQVQVLPQRLRIVRGTLFEDNYVALVLDNGDPLNNTTP 123
 QY 129 VTGASPGGLRELQRLSLTEILKGVLIQVLPQCYODTILWKDIFHKNQALTLIDTNR 188
 DB 124 -SSQAGLGLRELRLSLTEILKGVLIQVLPQCYODTILWKDIFHKNQALTLIDTNR 182
 QY 189 SRACHPCSPKGRGWGESSEDCQSLTRTVCAAGGCARCKGFLPTDCCHQCAAGCTGPK 248
 DB 183 N--CPRCSSACKSGGCGWGXQDCQTLTSVNCSSGCRCKGPKPSDCHVQCAAGCTGPK 240
 QY 249 HSDCLACHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPYNYLSTDV 308
 DB 241 DSDCLACHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPYNYLSTDV 300
 QY 309 GSCTLVCPHNEQVTA--EDG--TQRCCKSKPCARGTHSL 345
 DB 301 -ACTWCPKANKEVISVEPDQGTQCKEKGCEPKVCYGL 340

RESULT 6
 Q90836 PRELIMINARY; PRT; 527 AA.
 AC Q90836;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE EGF/TGF-alpha receptor precursor.
 GN C-ERBB.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92123214; PubMed=1732751;
 RA Flickinger T.W., Mailhe N.J., Kung H.-J.;
 RA "An alternatively processed mRNA from the avian c-erbB gene encodes a
 RT soluble, truncated form of the receptor that can block ligand-
 RT dependent transformation";
 RL Mol. Cell. Biol. 12:883-893(1992).
 DR EMBL; M77637; AAA48759.1; --
 DR PIR; A42032; A42032.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

DR GO:0007169; P:transmembrane receptor protein tyrosine kin. . . ; IEA.
 DR InterPro: IPR000494; EGFR_L domain.
 DR InterPro: IPR006211; Furin-like.
 DR InterPro: IPR006212; Furin repeat.
 DR InterPro: IPR009030; Grow_Fac_recep.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR SMART: SM00261; FU; 3.
 DR SMART: SM00261; FU; 3.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.
 KW Receptor; Signal.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 527 EGF/TGF-ALPHA RECEPTOR.
 SQ SEQUENCE 527 AA; 58353 MW; 764564ABCC095298 CRC64;
 Query Match 34.7%; Score 794; DB 13; Length 527;
 Best Local Similarity 46.0%; Pred. No. 1.4e-64;
 Matches 155; Conservative 55; Mismatches 109; Indels 18; Gaps 7;
 QY 11 LLLALLPGGAASST---OVCTGTDMLRLPASPTHLDMLRLHYOGCVQVGNLELTYL 65
 Db LLLALLPGGAASST---OVCTGTDMLRLPASPTHLDMLRLHYOGCVQVGNLELTYL 65
 QY 20 LLLALLGRVALCSAVEEKVKVCGTNNKLTQLGHVDFHTSLQRMNNCEVLSNLEITYV 79
 Db LLLALLGRVALCSAVEEKVKVCGTNNKLTQLGHVDFHTSLQRMNNCEVLSNLEITYV 79
 QY 66 PTNASLFLQDIQEVGYVLIHNOVRQVPLQRLIRIVRGTLQFEDNYALAVLDNGDPLNN 125
 Db PTNASLFLQDIQEVGYVLIHNOVRQVPLQRLIRIVRGTLQFEDNYALAVLDNGDPLNN 125
 QY 80 EHRDLTFLKTIQEVAGYVLIHNOVRQVPLQRLIRIVRGTLQFEDNYALAVLDNGDPLNN 138
 Db EHRDLTFLKTIQEVAGYVLIHNOVRQVPLQRLIRIVRGTLQFEDNYALAVLDNGDPLNN 138
 QY 126 TTPVTGASGGLBELQLRLSLTEILKGGVLIQNPOLCYQDTILWKDIFHNNQLALTLD 185
 Db TTPVTGASGGLBELQLRLSLTEILKGGVLIQNPOLCYQDTILWKDIFHNNQLALTLD 185
 QY 139 TQ-----GLRELPMKLSILNGGVKISNNPKLCNMDTVLWDDIITSRK-PLTVLD 189
 Db TQ-----GLRELPMKLSILNGGVKISNNPKLCNMDTVLWDDIITSRK-PLTVLD 189
 QY 186 -NRSRACHPCSPMKGSRGCSWESSEDQSLTRTVTCAGGCA-RCKGPLETDCHEQCAAG 243
 Db -NRSRACHPCSPMKGSRGCSWESSEDQSLTRTVTCAGGCA-RCKGPLETDCHEQCAAG 243
 QY 190 FASNLSSCPKCHPNTCTDHCWAGEGNCQTLKVIQAQCSGRGKVPDSCCHNQCAAG 249
 Db FASNLSSCPKCHPNTCTDHCWAGEGNCQTLKVIQAQCSGRGKVPDSCCHNQCAAG 249
 QY 244 CTGPKHSCLACLFHNSGICELHCPALVTNTDTFESMNPREGRYTFGASCVTACPVNY 303
 Db CTGPKHSCLACLFHNSGICELHCPALVTNTDTFESMNPREGRYTFGASCVTACPVNY 303
 QY 250 CTGPRESCLACRFKRDATCKDTPPLVLYNPTTYQMDVNPPEGKYSFGATCVRECPHY 309
 Db CTGPRESCLACRFKRDATCKDTPPLVLYNPTTYQMDVNPPEGKYSFGATCVRECPHY 309
 QY 304 LSTDVGSCTLVCPHLNHOEVTAEQGTQRCCKSPKCAR 340
 Db LSTDVGSCTLVCPHLNHOEVTAEQGTQRCCKSPKCAR 340
 QY 310 VVDHGSCVRSGNTDYEV-EENGVRKCKCKDGLCSR 345
 Db VVDHGSCVRSGNTDYEV-EENGVRKCKCKDGLCSR 345

RESULT 7
 QSERV6 PRELIMINARY; PRT; 643 AA.
 AC QSERV6;
 DT 01-MAR-2002 (TREMBLrel. 16, Created)
 DT 01-MAR-2002 (TREMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Epidermal growth factor receptor isoform 2.
 GN EGFR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/101, 129/SvJ, and 129/SvEvTAC;
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
 RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
 RA Balasubramanian S., Crossley T.O., Magnuson T.R., James C.D.,
 RA Mahle N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and mouse alternative Egr transcripts encoding truncated receptor isoforms."
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF275366; AG28046.1;
 DR EMBL; AF275366; AG28046.1; JOINED.
 DR EMBL; AF275365; AG28046.1; JOINED.
 DR MGD; MGI:95294; Egr.
 DR GO:0030139; C:endocytic vesicle; IDA.
 DR GO:0005622; C:intracellular; IDA.
 DR GO:0005515; F:protein binding; IPI.

DR InterPro: IPR000345; CytC_heme BS.
 DR InterPro: IPR000494; EGFR_L domain.
 DR InterPro: IPR006211; Furin-like.
 DR InterPro: IPR006212; Furin repeat.
 DR InterPro: IPR009030; Grow_Fac_recep.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR SMART: SM00261; FU; 3.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.
 KW Receptor.
 SQ SEQUENCE 643 AA; 71476 MW; DEF22002C8491131 CRC64;
 Query Match 34.5%; Score 789; DB 11; Length 643;
 Best Local Similarity 46.3%; Pred. No. 5.4e-64;
 Matches 156; Conservative 43; Mismatches 116; Indels 22; Gaps 6;
 QY 11 LLLALLPGGAASST---OVCTGTDMLRLPASPTHLDMLRLHYOGCVQVGNLELTYLPTN 68
 Db LLLALLPGGAASST---OVCTGTDMLRLPASPTHLDMLRLHYOGCVQVGNLELTYLPTN 68
 QY 14 LLLALLCAAGGALEKKVKVCGTNNKLTQLGHVDFHTSLQRMNNCEVLSNLEITYVQRN 73
 Db LLLALLCAAGGALEKKVKVCGTNNKLTQLGHVDFHTSLQRMNNCEVLSNLEITYVQRN 73
 QY 69 ASLSFLQDIQEVGYVLIHNOVRQVPLQRLIRIVRGTLQFEDNYALAVLDNGDPLNNTP 128
 Db ASLSFLQDIQEVGYVLIHNOVRQVPLQRLIRIVRGTLQFEDNYALAVLDNGDPLNNTP 128
 QY 74 YDLSFLKTIQEVAGYVLIHNOVRQVPLQRLIRIVRGTLQFEDNYALAVLDNGDPLNNTP 124
 Db YDLSFLKTIQEVAGYVLIHNOVRQVPLQRLIRIVRGTLQFEDNYALAVLDNGDPLNNTP 124
 QY 129 VTGASPGGLBELQLRLSLTEILKGGVLIQNPOLCYQDTILWKDI---FHKNNQLALTLD 184
 Db VTGASPGGLBELQLRLSLTEILKGGVLIQNPOLCYQDTILWKDI---FHKNNQLALTLD 184
 QY 125 -YGNRTGLRELPMKLSILNGGVKISNNPKLCNMDTVLWDDIITSRK-PLTVLD 180
 Db -YGNRTGLRELPMKLSILNGGVKISNNPKLCNMDTVLWDDIITSRK-PLTVLD 180
 QY 185 DTRSRACHPCSPMKGSRGCSWESSEDQSLTRTVTCAGGCA-RCKGPLETDCHEQCAAG 243
 Db DTRSRACHPCSPMKGSRGCSWESSEDQSLTRTVTCAGGCA-RCKGPLETDCHEQCAAG 243
 QY 181 -QSHSSCPKCHPNTCTDHCWAGEGNCQTLKVIQAQCSGRGKVPDSCCHNQCAAG 239
 Db -QSHSSCPKCHPNTCTDHCWAGEGNCQTLKVIQAQCSGRGKVPDSCCHNQCAAG 239
 QY 244 CTGPKHSCLACLFHNSGICELHCPALVTNTDTFESMNPREGRYTFGASCVTACPVNY 303
 Db CTGPKHSCLACLFHNSGICELHCPALVTNTDTFESMNPREGRYTFGASCVTACPVNY 303
 QY 240 CTGPRESCLACRFKRDATCKDTPPLVLYNPTTYQMDVNPPEGKYSFGATCVRECPHY 299
 Db CTGPRESCLACRFKRDATCKDTPPLVLYNPTTYQMDVNPPEGKYSFGATCVRECPHY 299
 QY 304 LSTDVGSCTLVCPHLNHOEVTAEQGTQRCCKSPKCAR 340
 Db LSTDVGSCTLVCPHLNHOEVTAEQGTQRCCKSPKCAR 340
 QY 300 VVDHGSCVRSGNTDYEV-EEDGIRKCKCKDGLCSR 335
 Db VVDHGSCVRSGNTDYEV-EEDGIRKCKCKDGLCSR 335

RESULT 8

QSERV5 PRELIMINARY; PRT; 655 AA.
 AC QSERV5;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Epidermal growth factor receptor (Epidermal growth factor receptor isoform 3).
 GN EGFR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RA Reiter J.L., Threadgill D.W., Danielson A.J., Schell C.,
 RA Lampland A.L., Balasubramanian S., Crossley T.O., Magnuson T.R.,
 RA Mahle N.J.;
 RT "Alternative Transcripts from the Human and Mouse EGFR Genes Encode Carboxy-Terminal Truncated Receptors."
 RT Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/101, 129/SvJ, and 129/SvEvTAC;
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
 RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
 RA Balasubramanian S., Crossley T.O., Magnuson T.R., James C.D.,
 RA Mahle N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and mouse alternative Egr transcripts encoding truncated receptor isoforms."
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF275366; AG28046.1;
 DR EMBL; AF275366; AG28046.1; JOINED.
 DR EMBL; AF275365; AG28046.1; JOINED.
 DR MGD; MGI:95294; Egr.
 DR GO:0030139; C:endocytic vesicle; IDA.
 DR GO:0005622; C:intracellular; IDA.
 DR GO:0005515; F:protein binding; IPI.

SQ SEQUENCE 1210 AA; 134840 MW; 62CD021C9DB32E18 CRC64;
Query Match 34.5%; Score 789; DB 11; Length 1210;
Best Local Similarity 46.3%; Pred. No. 1.2e-63;
Matches 156; Conservative 43; Mismatches 116; Indels 22; Gaps 6;
QY 11 LLALLPPCAA--STOVCTGTDMKRLPASPTHLDMLRLHLYQGVVQGNLELYLPTN 68
DB 14 LUTALCAAGALAEKKVCGTSENRLTQGTFFDHFLSLQRMNNECVLGNLEITYVQRN 73
QY 69 ASLSFLQDIQVQGVYLIHNRQVPLQRLRIVRGTLQFEDNYALAVLDNGDPLNTP 128
DB 74 YDLSFLKTIQVAGYVLIATNVTREPLENLIQIRGNALYENTYALAILSN----- 124
QY 129 VTGASPGGLRELQRLSLTEILKGGVLIQNPOLCVQDTILWKDI-----PHKNQALATLI 184
DB 125 -YGNRTGLRELPHENLQILICAVRFSNNPILCNWDTTQWRIIVQVPMNSMDL--- 180
QY 185 DYNRGRACHPCSPMCKSGRCWGESSEDCQSLRTVTCAGGCA-RCXGPIFTDCHEQCAAG 243
DB 181 -QSHPPSCPKDPCSPNCGWGGEECKLKIICAQCCSHRCGRSPSDCHNQCAAG 239
QY 244 CTGPKHSDCLACLFHNSGICELHCPALTYNTDTFESMPNPEGRYTFGASCVTACPNY 303
DB 240 CTGPRESCLVCKQFQDEATCKDTCPPLMLNPTTYQMDVNPFGKYSGATCVKCKPRNY 299
QY 304 LSTDVGSCTLVCPLENQVTAEDGTQRCCKSKPCAR 340
DB 300 VTDRGSCVRACPDYIEV-EEDGIRKCKKCDGPGCK 335
RESULT 10
Q8MIL8 PRELIMINARY; PRT; 1209 AA.
ID Q8MIL8
AC Q8MIL8
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Epidermal growth factor receptor.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim J.G., Vallet J.L., Nonneman D., Christenson R.K.;
RT "Characterization of uterine epidermal growth factor receptor
expression during the estrous cycle and early pregnancy in pigs."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY117054; AM77472.1; -
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0005489; F:electron transporter activity; IEA.
DR GO: GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR InterPro: IPR000345; CytC heme BS.
DR InterPro: IPR000494; EGFR_L domain.
DR InterPro: IPR006211; Purin-like.
DR InterPro: IPR009030; Grow_fac_recep.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR008266; Tyr_kinase_AS.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF00069; Pkinase; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 1.

DR SMART; SMO0220; S_TKC; 1.
DR SMART; SMO0219; TYRK; 1.
DR PROSITE; PS00190; CITOCHROME_C; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
KW ATP-binding; Receptor; Transferase.
SQ SEQUENCE 1209 AA; 133531 MW; 268E3FB1E36FP90F CRC64;
Query Match 34.5%; Score 788; DB 6; Length 1209;
Best Local Similarity 44.6%; Pred. No. 1.5e-63;
Matches 150; Conservative 55; Mismatches 113; Indels 18; Gaps 4;
QY 12 LLALL-----PPGAASSTOVCTGTDMKRLPASPTHLDMLRLHLYQGVVQGNLELYL 65
DB 11 LLALLAHFQSPALAEKKVCGTSENRLTQGTFFDHFLSLQRMNNECVLGNLEITYM 70
QY 66 PTNASLFLQDIQVQGVYLIHNRQVPLQRLRIVRGTLQFEDNYALAVLDNGDPLN 125
DB 71 QNSVLSFLKTIQVAGYVLIATNVTREPLENLIQIRGNALYENTYALAILSN----- 124
QY 126 TTPVTGASPGGLRELQRLSLTEILKGGVLIQNPOLCVQDTILWKDIHKNQALATLI 185
DB 125 -YGNRTGLRELPHENLQILICAVRFSNNPILCNWDTTQWRIIVQVPMNSMDL--- 180
QY 186 DYNRGRACHPCSPMCKSGRCWGESSEDCQSLRTVTCAGGCA-RCXGPIFTDCHEQCAAG 244
DB 181 -QSHPPSCPKDPCSPNCGWGGEECKLKIICAQCCSHRCGRSPSDCHNQCAAG 240
QY 245 CTGPKHSDCLACLFHNSGICELHCPALTYNTDTFESMPNPEGRYTFGASCVTACPNY 304
DB 241 CTGPRESCLVCKQFQDEATCKDTCPPLMLNPTTYQMDVNPFGKYSGATCVKCKPRNY 300
QY 305 STDVGSCTLVCPLENQVTAEDGTQRCCKSKPCAR 340
DB 301 VTDRGSCVRACSDSYEV-EEDGVRKCKKCDGPGCK 335
RESULT 11
Q9QX70 PRELIMINARY; PRT; 1209 AA.
ID Q9QX70
AC Q9QX70
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Epidermal growth factor receptor.
GN EGFR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Fisher; TISSUE=Liver;
RX MEDLINE=9025888; PubMed=2342466;
RA Patch L.A., Harris J., Raymond V.W., Blasband A.J., Lee D.C.,
RA Earp H.S.;
RT "A truncated, secreted form of the epidermal growth factor receptor is
encoded by an alternatively spliced transcript in normal rat tissue."
RL Mol. Cell. Biol. 10:2973-2982(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=Fisher; TISSUE=Liver;
RA Patch L.A.;
RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN=Fisher; TISSUE=Liver;
RA Guttridge K., Dawson T.L., Earp H.S.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: M37394; AAF14008.1; -
DR HSP; P11362; 1PKK.
DR GO: GO:0016020; C:membrane; IEA.

DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO: GO:0004872; P:receptor activity; IEA.
DR GO: GO:0016740; P:transferase activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. .; IEA.
DR InterPro: IPR000494; EGFR_L domain.
DR InterPro: IPR006211; Furin-like.
DR InterPro: IPR006212; Furin repeat.
DR InterPro: IPR009030; Grow fac recep.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR008266; Tyr_kinase_AS.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF01030; Recep_L domain; 2.
DR PRINTS: PR0109; TYRKINASE.
DR PRODOM: PD000001; Prot_kinase; 1.
DR SMART: SM00261; FU; 3.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Kinase; Receptor; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 1209 AA; 134891 MW; 96FE7F6CC1B7773 CRC64;

Query Match 34.2%; Score 781.5; DB 11; Length 1209;
Best Local Similarity 46.0%; Pred. No. 66-63;
Matches 157; Conservative 44; Mismatches 117; Indels 23; Gaps 7;
QY 3 LAALCRWGLLALLPPGA--ASTQVCTGTDMKRLPASPETHLMLRHLHYOGQVVGNGLE 61
DB 15 LAALCAAG-----GALEKKVCOGT-SNRLTQGTGFEDHFLSLQRMNCEVVLGNLE 66
QY 62 LTYLPTNASLFDIOIQVQGVYLIAHNOVQVPLQRLIRVGTOLFDENVALAVLNGD 121
DB 67 ITYQRYNDLSFLKTIQVAGVYLIATNVERIPLENQIIRGNALYENTYVALVLSN-- 124
QY 122 PLNNTPTVGTASPGRLQLRLSLTEILKGGVLIQRNPOLCYQDTILMKDIFHKNLAL 181
DB 125 -----YGNKTGLRLPKNLQELILGAVRFSNNPILCMETIQWRDIV-QDVLSN 175
QY 182 TLIDNRS-RACHPCSPCKSRGWGESSEDCOSLTRVTCAGGCA-RCKGPLPTDCCHQ 239
DB 176 MSMDVQRHLTGCPKDPSPCNWSCWGRGEENCQKLTIKIACQCSRRCRGSPSDCCHNQ 235
QY 240 CAAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMENPEGRYTFGASCVTAC 299
DB 236 CAAGCTGPRSDCLVCHFRDEATCKDTCPPLMLYNTTYQMDVNPESGYSFGATCVKKC 295
QY 300 PYNVLTDSVGSCTLVCPLNHQEVTAEDGTQRCCKSPCAR 340
DB 296 PRNVVVDHSGSVRACGPDYEV-EEDGVSKCKKCDGCPCK 335

RESULT 12
Q9ESE0 PRELIMINARY; PRT; 478 AA.
AC Q9ESE0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Epidermal growth factor receptor related protein.
GN ERAP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=gastro-duodenal mucous;
RA Yu Y., Moshier J.A., Majumdar A.P.N.;
RT "Cloning of a novel EGFR-related peptide: A putative negative

regulator of EGFR.",
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF187818; AAG17037.2; -
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO: GO:0004872; P:receptor activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. .; IEA.
DR InterPro: IPR000494; EGFR_L domain.
DR InterPro: IPR006211; Furin-like.
DR InterPro: IPR009030; Grow_fac_recep.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF01030; Recep_L domain; 2.
KW Receptor.
SQ SEQUENCE 478 AA; 53233 MW; CF873A8376C519E5 CRC64;

Query Match 33.8%; Score 773; DB 11; Length 478;
Best Local Similarity 45.7%; Pred. No. 11e-62;
Matches 154; Conservative 43; Mismatches 118; Indels 22; Gaps 6;
QY 11 LLALLPPGAA--STQVCTGTDMKRLPASPETHLMLRHLHYOGQVVGNGLEITYPTN 68
DB 14 LLTALCAAGALBEKKVCOGT-SNRLTQGTGFEDHFLSLQRMNCEVVLGNLEITYVQRN 73
QY 69 ASLSPLDIOEVQGVYLIAHNOVQVPLQRLIRVGTOLFDENVALAVLNGDPLNNTTP 128
DB 74 YDSLFLKTIQVAGVYLIATNVERIPSEDLQIRGNALYENTYVALVLSN----- 124
QY 129 VTGASPGRLQLRLSLTEILKGGVLIQRNPOLCYQDTILMKDI----FKNNQLALTLI 184
DB 125 -YGNKTGLRLPKNLQELILGAVRFSNNPILCMETIQWRDIVQVNFMSNMDL--- 180
QY 185 DYNRSRACHPCSPCKSRGWGESSEDCOSLTRVTCAGGCA-RCKGPLPTDCCHQCAAG 243
DB 181 -QSHPSCKPKDPSPCNWSCWGRGEENCQKLTIKIACQCSRRCRGSPSDCCHNQCAAG 239
QY 244 CTGPKGSDCLACLHFNHSGICELHCPALVTYNTDTFESMENPEGRYTFGASCVTACPNY 303
DB 240 CTGPKGSDCLVCHFRDEATCKDTCPPLMLYNTTYQMDVNPESGYSFGATCVKNCPRN 299
QY 304 LSTDVGSCTLVCPLNHQEVTAEDGTQRCCKSPCAR 340
DB 300 VTDHSGSVRACGPDYEV-EEDGIRKCKKCDGCPCK 335

RESULT 13
Q7SZF7 PRELIMINARY; PRT; 1191 AA.
AC Q7SZF7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Epidermal growth factor receptor.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22797244; PubMed=12915231;
RA Goishi K., Lee P., Davidson A.J., Nishi E., Zon L.I., Klagsbrun M.;
RT "Inhibition of zebrafish epidermal growth factor receptor activity results in cardiovascular defects.";
RL Mech. Dev. 120:811-822(2003).
RL EMBL: AY154658; AAM47184.1; -.
KW Receptor.
SQ SEQUENCE 1191 AA; 132483 MW; F4A7472D60A6860A CRC64;

Query Match 33.1%; Score 757.5; DB 13; Length 1191;
Best Local Similarity 43.5%; Pred. No. 9.6e-61;
Matches 151; Conservative 52; Mismatches 129; Indels 15; Gaps 5;

```

QY 10 GLLIALLPPGAA-----ASTQVCTGTDMLRLPASPETHLMDLRLHLYQCCQVQGNLELTLYLPT 65
Db 11 GLLFSLARGSEVGNQAVCPGTENGLSVTDAENQYQTLKLYEYECRVWNGNLEIVLTGH 70
QY 66 PTNASFQDIOQEVQGYVLIHNOVRQVPLQRLIRIVRGTPQLFEDNYALAVLDNGDPLNNT 125
Db 68 TEKYDLSFLKSIQEVQGYVLIHNOVRQVPLQRLIRIVRGTPQLFEDNYALAVLDNGDPLNNT 124
QY 126 TPTVTCASPGGLRELQRLSRLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQALATLIDN 185
Db 125 SI-----BQVKELPLTSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQALATLIDN 178
QY 186 TNRSRACHPCSPMKGRSGWSSBDCQSLTRTVCAAGCA-RCKGPLPTDCCHQCAAGC 244
Db 179 ISVKNCKGKDCSCFNGSCWGTGPKQKQKTKVICAQCSGCKGPRPIDCCNEHCAAGC 238
QY 245 TQPKSDCLACHFNHSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACPNYLT 304
Db 239 TQPRPTDCLACKDFQDEGTCKDACPRMLYDPTNTHQALPAPYKYSFGATCIKTCFPHNV 298
QY 305 STDVGSCTLVCLPILHNOEVTAEQGTQCEKCKSPCAR 340
Db 299 VTDHGACVTCSPGTTEYVD-EGSVKCKRCEGLCPKVCNGLGKGPJA 344

```

RESULT 14

```

Q9BUD7 PRELIMINARY; PRT; 331 AA.
AC Q9BUD7
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Similar to v-erb-b2 avian erythroblastic leukemia viral oncogene
DE homolog 3 (v-erb-b2 erythroblastic leukemia viral oncogene homolog 3)
DE (Avian)
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kalcine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length CDSS in BD Creator (TM) System Donor
RT vector";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002706; AAH02706.1; -
DR EMBL; BT007226; AAP35890.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR InterPro; IPR000494; EGFR_L domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR Pfam; PF00757; Furin-Like; 1.
DR SMART; SM00261; FU; 2.
SQ SEQUENCE 331 AA; 36489 MW; 45B8EBEE683FE7B8 CRC64;

```

```

Query Match 32.1%; Score 734; DB 4; Length 331;
Best Local Similarity 44.0%; Pred. No. 2.8e-59;
Matches 147; Conservative 49; Mismatches 120; Indels 18; Gaps 8;

```

```

QY 10 GLLIALLPPGAA-----STQVCTGTDMLRLPASPETHLMDLRLHLYQCCQVQGNLELTLYLPT 67
Db 11 GLLFSLARGSEVGNQAVCPGTENGLSVTDAENQYQTLKLYEYECRVWNGNLEIVLTGH 70
QY 68 NAFSLFQDIOQEVQGYVLIHNOVRQVPLQRLIRIVRGTPQLFEDNYALAVLDNGDPLNNT 127
Db 71 NADLSFLQIWEIVTGYVLIHNOVRQVPLQRLIRIVRGTPQLFEDNYALAVLDNGDPLNNT 125
QY 128 PVTGASPGGLRELQRLSRLTEILKGGVLIQRPOLCYQDTILWKDIFHKNQALATLIDN 187
Db 126 ----NSSHALRQLRLTQLTEILSGGVYIEKNDKLCMDTIDWRDIVRDRD---ABIVVKD 178
QY 188 RSRACHPCSPMKGRSGWSSBDCQSLTRTVCAAGC-ARCKGPLPTDCCHQCAAGC 246
Db 179 NCRSPCPCHVECKG-RCWGPGSDEQTLTKTKICAPQNGHCFGPNPQCHDECAAGC 237
QY 247 PKHSDCLACHFNHSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACPNYLT 306
Db 238 QDPTDCAFRHFDGACVPRCPOLVYNKLTFOLEPNHTKYQYGGVCAVSCPHNV-V 296
QY 307 DVGSGCTLVCLPILHNOEVTAEQGTQCEKCKSPCAR 340
Db 297 DQTSVRACPPTDQMEVD-KNGLKMCPCGGLCPK 329

```

RESULT 15

```

Q9BG66 PRELIMINARY; PRT; 149 AA.
AC Q9BG66
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Receptor tyrosine kinase ErbB2 (Fragment)
OS Oryctolagus cuniculus (Rabbit)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Kuechenhoff A.,
RA Tetens F., Fischer B.;
RT "ErbB genes and epidermal growth factor- (EGF-) like ligands in the
RT peri-implantation rabbit uterus and blastocyst.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF333178; AAK14371.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004714; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow_fac_recep.
DR Pfam; PF00757; Furin-like; 1.
DR SMART; SM00261; FU; 2.
DR KINASE.
FT NON TER 1 1
FT NON TER 149 149
SQ SEQUENCE 149 AA; 16240 MW; 7CB3792A54FC49BA CRC64;

```

```

Query Match 31.6%; Score 723; DB 6; Length 149;
Best Local Similarity 81.2%; Pred. No. 1e-58;
Matches 121; Conservative 8; Mismatches 20; Indels 0; Gaps 0;

```

```

QY 158 NPQCYQDTILWKDIFHKNQALATLIDNRSRACHPCSPMKGRSGWSSBDCQSLTR 217
Db 1 NPQCYQDTILWKDIFHKNQALATLIDNRSRACHPCSPMKGRSGWSSBDCQSLTR 60
QY 218 TVCAGGACRCKGPLPTDCCHQCAAGCTGPGKSDCLACHFNHSGICELHCPALVTYNTD 277
Db 61 TICAGGACRCKGQLPTDCCHQCAAGCTGPGKSDCLACHFNHSGICELHCPALVTYNTD 120

```

Qy	278	TRESMENPEGRYTFGASCVTACPNYLST	306
Dd	121	TFESMNPGRYTFGASCVTTCFYNLST	149
RESULT 16			
QSAW81	QSAW81	PRELIMINARY;	PRT; 1305 AA.
ID	Q8AW81	Q8AW81	
AC	Q8AW81	Q8AW81	
DT	01-MAR-2003	(TrEMBLrel. 23, Created)	
DT	01-OCT-2003	(TrEMBLrel. 23, Last sequence update)	
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)	
DE	SI:d2150112.1	(Novel protein similar to ErbB (v-erb-b erythroblastic leukemia viral oncogene homolog, neuro/glioblastoma derived oncogene homolog))	
DE	leukemia	viral oncogene homolog, neuro/glioblastoma derived oncogene homolog	
GN	SI:D2150112.1		
OS	Brachydanio rerio (Zebrafish) (Danio rerio)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;		
OC	Cyprinidae; Danio		
OX	NCB1_TaxID=7955;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Schra H.;		
EL	Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL; AL591365; CAD58760.1; --		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.		
DR	GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.		
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.		
DR	GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . ; IEA.		
DR	InterPro; IPR000494; EGFR_L domain.		
DR	InterPro; IPR006211; Furin-like.		
DR	InterPro; IPR009030; Grow_fac_recep.		
DR	InterPro; IPR000719; Prot_kinase.		
DR	InterPro; IPR000719; Prot_kinase.		
DR	InterPro; IPR002230; Ser_thr_pkinase.		
DR	InterPro; IPR001245; Tyr_pkinase.		
DR	pfam; PF00757; Furin-like; 1.		
DR	pfam; PF01030; pkinase; 1.		
DR	PRINTS; PR00109; TYRKINASE.		
DR	PRODOM; PD000001; Prot_kinase; 1.		
DR	SMART; SM00220; S_TKc; 1.		
DR	SMART; SM00219; Tyrc; 1.		
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.		
DR	PROSITE; PS00109; PROTEIN KINASE DOM; 1.		
DR	PROSITE; PS00111; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00112; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00113; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00114; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00115; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00116; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00117; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00118; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00119; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00120; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00121; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00122; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00123; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00124; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00125; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00126; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00127; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00128; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00129; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00130; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00131; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00132; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00133; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00134; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00135; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00136; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00137; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00138; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00139; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00140; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00141; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00142; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00143; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00144; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00145; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00146; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00147; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00148; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00149; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00150; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00151; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00152; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00153; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00154; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00155; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00156; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00157; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00158; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00159; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00160; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00161; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00162; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00163; PROTEIN KINASE TYP; 1.		
DR	PROSITE; PS00164; PROTEIN KINASE TYP;		

QY 56 VQGNLELTPTNLSLQDIOEVOGYVLIHNRQVPLQRLRIVRGTQTFEDNYALA 115
DB 57 VLENLEITVQGNLSEFQSGQVEGVYVLIHNRQVPLQRLRIVRGTQTFEDNYALA 116
QY 116 VLDNGDPLNNTPTVTPGASGGRLRLQLRLSLRLKGGVLIQNPOLCYQDTLLWKDIFHK 175
DB 117 VMSYQK-NPSSP--DVYQVGLKQLSLNLTILSGGVKSHNPLLCNVETINWVDIVDK 173
QY 176 NNQLALTLDITNRSRACHPCSPMKSGSRGSGSSDCCSLRTVTCAGGC-ARCKGPLPTD 234
DB 174 TSNPTNLIPIHAFERQCKDQKDCGVNCSWAPGCHQKFTKLLCAEQNRCRCRCPKPID 233
QY 235 CCHQCAAGCTGPKHSDGLACLFHNSGICELRCPALVTYNTDTFESMNPENRGTFGAS 294
DB 234 CCNEHCAGGCTGPRATDCLACRDFNDGFCDKTCTPPKIYDIVSHQVVDNPNKITYFGAA 293
QY 295 CVTACPNYSLDVSCTLVCPHNRQVTAEDGTORCEKCKPCAR 340
DB 294 CVKECPNSYVWTE-GACVRSAGMLEVD-ENGKRSCKPCDGVCPK 337
RESULT 18
P79754 PRELIMINARY; PRT; 1328 AA.
ID P79754;
AC P79754;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Erbb3.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthopterygii; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
R2 SEQUENCE FROM N.A.
RX MEDLINE=99177347; PubMed=10077531;
RA Gelliner K., Brenner S.;
RT "Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu
rubripes";
RL Genome Res. 9:251-258(1999).
DR EMBL; AF056116; AAC34391.1; --
DR HSSP; P11362; 1FKC.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006458; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow fac recep.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00065; pkinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR PRINTS; PD00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 3.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; kinase; transferase; tyrosine-protein kinase.
SQ SEQUENCE 1328 AA; 148613 MW; A3330392585647E9 CRC64;
Query Match 29.0%; Score 662.5; DB 13; Length 1328;
Best local Similarity 42.1%; Pred. No. 6.5e-52;
Matches 143; Conservative 47; Mismatches 125; Indels 25; Gaps 11;

QY 9 WGLLALLAPP--GAASQ---VCTGTDMLRLPASPEHLDMLHLHYOGCVVQGNLEL 62
DB 4 WRLILMCVASLRASSQTQEAQVCTGNGUSSGTSQENQYNLNRKYKGEIINGNDEI 63
QY 63 TYLPTNASLSLQDIOEVOGYVLIHNRQVPLQRLRIVRGTQTFEDNYALAVLDNGDP 122
DB 64 TOIESNDFSELTIRETVGVYVLIHNRQVPLQRLRIVRGTQTFEDNYALAVLDNGDP 120
QY 123 LNNVTPTVTPGASGGRLRLQLRLSLRLKGGVLIQNPOLCYQDTLLWKDIFHKNNQLALT 182
DB 121 ----YPKDG--PSGLNQLGLMNLTEILDGGVQIINNKLRYGVPWYMRDII--RNNDAPIE 173
QY 183 LIDTNRSRACHPCSPMKSGSRGSGSSDCCSLRTVTCAGGC-ARCKGPLPTDCCHEOCA 241
DB 174 IQPNERGVCH--KSC-GNYCWPCKDQCOILKTVCAPOCNDRCFGTSRDCCHISCA 229
QY 242 AGCTGPKHSDGLACLFHNSGICELRCPALVTYNTDTFESMNPENRGTFGASCVTACPY 301
DB 230 AGCKGPLDTCFACRLFNDSGACYPQCPQTLLIYNKQTFQMETNPNRYQYSGICVSCQPT 289
QY 302 NYLSTDVSGCTLVCPHNRQVTAEDGTOR-CEKCKPCAR 340
DB 290 HFV-VDSGSCVSVCPDKMEV--ERGSQRCGLCGLCPK 326
RESULT 19
Q9PSH2 PRELIMINARY; PRT; 599 AA.
ID Q9PSH2;
AC Q9PSH2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Epidermal growth factor receptor (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
R2 SEQUENCE FROM N.A.
RX MEDLINE=94020816; PubMed=8414496;
RA Callaghan T., Antczak M., Flickinger T., Raines M., Myers M.,
Kung H.J.;
RT "A complete description of the EGF-receptor exon structure:
implication in oncogenic activation and domain evolution";
RL Oncogene 8:2939-2948(1993).
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0006458; P:epidermal growth factor receptor activity; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow fac recep.
DR Pfam; PF00757; Furin-like; 2.
DR Pfam; PF01030; Recep_L_domain; 2.
DR SMART; SM00261; FU; 3.
SQ SEQUENCE 599 AA; 66363 MW; FEAB46D293D991BD CRC64;
Query Match 28.3%; Score 647; DB 13; Length 599;
Best local Similarity 27.4%; Pred. No. 6.3e-51;
Matches 155; Conservative 52; Mismatches 111; Indels 248; Gaps 11;
QY 24 QVCTGTDMLRLPASPEHLDMLHLHYOGCVVQGNLELTPTNLSLQDIOEVOGY 83
DB 5 KVCQGNKNTQIGHVEDHTSLQRYNNCEVLSNLEIYVHNRLDLSFKTIQRYVAGY 64
QY 84 VLIHNRQVPLQRLRIVRGTQTFEDNYALAVLDNGDPPLNNTPTVTPGASGGRLRLQLR 143
DB 65 VLIHNRQVPLQRLRIVRGTQTFEDNYALAVLDNGDPPLNNTPTVTPGASGGRLRLQLR 115
QY 144 SLTEILKGGVLIQNPOLCYQDTLLWKDIFHKNNQLALTLD----- 185

Db 116 RLSEILNGVGIKSNPKLNMDFVLDIIDSRL-PLTVLDFASNLSSVEBNGVRCKK 174
Qy 186 -----
Db 175 CDGLCKVNGIGIGELKILSINATNIDSPQNKCTKINGDVSVILPVLGDAFTKTLPLD 234
Qy 186 -----
Db 235 PKKLDVFRVKEISGFLLIQWPDNATLYAFENLEIRGRTKOHQGYSLAVNLKIQSLG 294
Qy 186 -----
Db 295 LRSLEKEISGDIAIMKXNKLVCYADTMWRSLFATOSQKTKLIONKNDCKPKNCTE 354
Qy 202 SRCWGESSEDSQSLRTVWAGCA-RCKGPLPTDCHEQCAGCTGPKHSDCLACLHFNH 260
Db 355 DHCWAGBNQCNQTLTKVCAQCSGRGKVPSCDCHNQCAAGCTGPPRESCLACRKF 414
Qy 261 SGICELHCPALVTYNTDFESMPNPEGRVTFGASCVTACPNY-----LSTDVG-- 309
Db 415 DATCKDTCPPLVLYNFTYQMDVNPFGKYSFGATVRCPEHTADRHVCDPLCSVDGCGW 474
Qy 310 -----SCLVLC-----PLNQEVTAEDGTQRCCKSK 336
Db 475 PGPFHCFSCRPFSRQKCEVKQCNILQGEPRFEDSKCLPCHS-ECLVQNSTAVNTCSG 533
Qy 337 P-----CAR---GTHSLRPAAY 352
Db 534 PGPDHCKMAHFDGPHCVKACPAVG 559

RESULT 20

Q8MLW0
ID Q8MLW0 PRELIMINARY; PRT: 1377 AA.
AC Q8MLW0:
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE CG10079-PA.
GN EGFR OR CG10079.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Abril J.F., Adayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley B.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann M.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöckner A., Gong F., Gorrill J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Murzy D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleib J.M.,
Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,
Shue B.C., Siden-Klams I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
"The genome sequence of Drosophila melanogaster";
Science 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
Celinker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
Genzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
Ibegwan C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
Pacleib J., Paragass V., Park S., Patel S., Pfeiffer B.,
Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler P.,
Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
"Sequencing of Drosophila melanogaster genome";
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
Tupy J.L., Bergman C., Berman B., Carlson J.W., Celinker S.E.,
Clamp M., Drysdale R., Emmert D., Frise S., Ge Grey A., Harris N.,
Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
Searle S.M.J., Smith E., Shu S., Smitniak F., Whitfield E.,
Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
"Annotation of Drosophila melanogaster genome";
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
Adams M.D., Celinker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A.
FlyBase;
Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
EMBL; AE003454; AAM70919.1; -;
PIR; A27131; A27131.
FlyBase; FBgn0003731; Egfr.
GO; GO:0007469; P:antennal morphogenesis; NAS.
GO; GO:0006916; P:anti-apoptosis; NAS.
GO; GO:0030381; P:eggshell pattern formation (sensu Insecta); IGI.
GO; GO:0007456; P:eye morphogenesis (sensu Drosophila); IMP.
GO; GO:0007390; P:germ-band shortening; IMP.
GO; GO:0007449; P:imaginal disc development; IMP.
GO; GO:0007479; P:leg disc proximal/distal pattern formation; IMP.
GO; GO:0008071; P:maternal determination of dorsal/ventral ax.; IMP.
GO; GO:0007477; P:notum morphogenesis; IMP.
GO; GO:0007314; P:ooocyte anterior/posterior axis determination; NAS.
GO; GO:0045468; P:regulation of R8 spacing; NAS.
GO; GO:0016300; P:regulation of R8 spacing; NAS.
GO; GO:0007476; P:wing morphogenesis; IMP.
InterPro; IPR000345; CytC_heme_BS.
InterPro; IPR000494; EGFR_L_domain.
InterPro; IPR006211; Furin-like.
InterPro; IPR006212; Furin repeat.
InterPro; IPR009030; Grow_fac_recep.
InterPro; IPR000719; Prot_kinase.

```
DR InterPro: IPR002290; Ser thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR InterPro: IPR008266; Tyr_pkinase_AS.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00261; FU; 7.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TyKc; 1.
DR PROSITE: PS00190; CYTOCHROME_C_3.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Transferase.
SQ SEQUENCE 1377 AA; 153601 MW; 38AFA4A9D5C954 CRC64;

Query Match 26.3%; Score 601; DB 5; Length 1377;
Best Local Similarity 35.8%; Pred. No. 3.2e-46;
Matches 127; Conservative 47; Mismatches 135; Indels 46; Gaps 8;

QY 9 WGLLLALLPPGAAST-----QVCTGTDMLRLPASPETHLDMLRLHYQGC 53
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 21 WSVLLILACMASITSSVSNAGYVDNGMKVICIGTSRLSVPSNKEHHYRNLDRYTNC 80
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 54 QVQGNLELYLPT-NASLSFLQDIQVQGVLIHNVQVPLQRLIVRGTLQF----- 108
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 81 TYVDGNLELTLWPNENLDLSFLDNIREVTGYTILSHVDVKKVFPFKLQIIRGTLFSLV 140
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 109 -EDNYALAVLDNGDPLNNTPTVTGASPGGLRELQSLTEILKGGVLIQNPOLCYQDTI 167
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 141 EEEKYALFV-----TYSKMYTLLEIPDLRDVLNGQGVFNHYNLCHMTI 184
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 169 LMKDIPKNNQALALTIDNRSRACHPCSPMKGRGSGESSEDQSILTRTVCAAGCA-- 225
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 185 QWSEIVSNGTDAYNYDFTAPERECPKCHESCTHG-CWGEKPNCKQKFSKLTCSQCAGG 243
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 226 RCKGPLEPTDCHEQCAAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTFESMNP 285
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 244 RCYKPKRECCHLFCAGGCTGTQKDCIACQFFBEGVCKECPMPKYNPTTYVLETPN 303
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 286 EGRYTFGASCTVACPYNLYSLDVGSCTLVCPLNHNEVTAEDGTQCEKSKPCAR 340
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 304 EGYKAYGATCVKECP-GHLLRDNGACVRSQDQKMDKGE-----CVPNGPCPK 352
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 21
Q86N22 PRELIMINARY; PRT; 1322 AA.
AC Q86N22;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE LP05058p (Fragment).
GN EGFR.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dreanek D., Farfan D., Frise B.,
RA George R., Gonzalez M., Guarin H., Koonmiller B., Li P., Liao G.,
RA Miranda A., Mungali C.J., Nunoo J., Pacled J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBAJ databases.
DR EMBL; BT003562; AAC03566.1; -.
DR GO; GO:0016020; C.membrane; IEA.
```

```
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005489; P:electron transporter activity; IEA.
DR GO; GO:0005006; P:epidermal growth factor receptor activity; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006488; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; IEA.
DR InterPro: IPR000345; CytC_heme_BS.
DR InterPro: IPR00494; EGFR_L_domain.
DR InterPro: IPR006211; Furin-like.
DR InterPro: IPR006212; Furin repeat.
DR InterPro: IPR003030; Grow_fac_recep.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR InterPro: IPR008266; Tyr_pkinase_AS.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00261; FU; 7.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TyKc; 1.
DR PROSITE: PS00190; CYTOCHROME_C_3.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
FT NON TER 1
SQ SEQUENCE 1322 AA; 147699 MW; E778CEA19BB313B4 CRC64;

Query Match 25.2%; Score 576.5; DB 5; Length 1322;
Best Local Similarity 37.2%; Pred. No. 5.6e-44;
Matches 119; Conservative 43; Mismatches 127; Indels 31; Gaps 7;

QY 29 TDMKLRIPASPETHLDMLRLHYQGVQVQGNLELYLPT-NASLSFLQDIQVQGVLIH 87
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 TKSRLSVPSNKEHHYRNLDRYTCTYVDGNLELTLWPNENLDLSFLDNIREVTGYTILS 60
QY 88 HNOYRQVPLQRLIVRGTLQF-----EDNYALAVLDNGDPLNNTPTVTGASPGGLRELQ 142
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 HVDKVKVFPFKLQIIRGTLFSLSVBEKYLAFV-----TYSKMYTLLEI 104
QY 143 RSLTEILKGGVLIQNPOLCYQDTILMKDIPKNNQALALTIDNRSRACHPCSPMKGS 202
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 PDLRDVLNGQGVFNHYNLCHMTIQWSEIVSNGTDAYNYDFTAPERECPKCHESCTHG 164
QY 203 RCWGESSEDQSILTRTVCAAGCA--RCKGPLEPTDCHEQCAAGCTGPKHSDCLACHFNH 260
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 165 -CWGEKPNCKQKFSKLTCSQCAGGRCYKPKRECCHLFCAGGCTGTQKDCIACNFPD 223
QY 261 SGICELHCPALVYNTDTFESMNPPEGRYTFGASCTVACPYNLYSLDVGSCTLVCPLNH 320
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 224 EGVCKECPMPKYNPTTYVLETPNPEKAYGATCVKECP-GHLLRDNGACVRSQDQKMDK 282
QY 321 EVTAEDGTQCEKSKPCAR 340
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 283 DKGE-----CVPNGPCPK 297

RESULT 22
Q9BIH9 PRELIMINARY; PRT; 1433 AA.
AC Q9BIH9
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative epidermal growth factor receptor (Fragment).
GN EGFR.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=7165;
```

```

3N  AC Q9W6F6;
RP  SEQUENCE FROM N.A.
RC  STRAIN-SUA;
RA  Lycett G.O.;
RT  "Cloning, expression and localisation of the Anopheles gambiae
RL  epidermal growth factor receptor.";
DR  EMBL; AJ301655; CAC35008.1; -.
DR  HSSP; P11362; 1FGK.
DR  GO; GO:0016020; C:membrane; IEA.
DR  GO; GO:0005524; F:ATP binding; IEA.
DR  GO; GO:0005489; F:electron transporter activity; IEA.
DR  GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR  GO; GO:0004872; F:receptor activity; IEA.
DR  GO; GO:0016740; F:transferase activity; IEA.
DR  GO; GO:0006118; P:electron transport; IEA.
DR  GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR  GO; GO:0007189; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR  InterPro; IPR000345; CYC_heme_BS.
DR  InterPro; IPR000494; EGFR_L_domain.
DR  InterPro; IPR006211; Furin-like.
DR  InterPro; IPR009030; Grow_fac_recep.
DR  InterPro; IPR000719; Prot_kinase.
DR  InterPro; IPR001245; Tyr_kinase.
DR  InterPro; IPR008266; Tyr_kinase_AS.
DR  Pfam; PF00757; Furin-like; 1.
DR  Pfam; PF00069; pkinase; 1.
DR  PRINTS; PS00109; TYRKINASE.
DR  ProDom; PD000001; Prot_kinase; 1.
DR  SMART; SM00261; FU; 7.
DR  SMART; SM00219; TyrcK; 1.
DR  PROSITE; PS00190; CYTOCHROME C; 4.
DR  PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR  PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR  PROSITE; PS00109; PROTEIN KINASE_TYR; 1.
KW  ATP-binding; Kinase; Receptor; Transferase; Tyrosine-protein kinase.
FT  NON TER 1
SQ  SEQUENCE 1433 AA; 19585 MW; E3D9D89967724F07 CRC64;

Query Match 25.0%; Score 571.5; DB 5; Length 1433;
Best Local Similarity 37.2%; Pred. No. 1.8e-43;
Matches 120; Conservative 47; Mismatches 125; Indels 31; Gaps 7;

QY 26 CTGTDMLRLASPEHLDMRLHYQGVQVQGNLELYLPNLSLPLQDIQVQGVVL 85
DB 1 CIGTNGSVSPANREHYKLNDRYNTCTYVDGNELEITWQITDLNPLQIREVTGVL 60
QY 86 IAHNQVRQVPLQRLRIVRGTLF-----EDYALAVLNGDPLNNTPTVTGASPGGLREL 140
DB 61 ISLYDLPLQVILPRELQIRGRTTFKLNKWEAYGLFV-----SPSHMNTL 104
QY 141 QRLSITELKGGVLIQRPOLCYQDTILWKDI-FHKNQLALTLDITNRSACHPCSPMC 199
DB 105 ELPALRDLTGGVSGFFNNYNLCHMKSNKWEILLAPQTSMQYTFNFSPEVCPCHPSC 164
QY 200 XGSRGWSSSDQSLRTVCAGGCA--RCKGPLPTDCCHQCAGCTGPKHSDCLACLH 257
DB 165 EVG-CWGSAGNCRQFSLKNSPQCSQGRCFGPKPRECHLFAGGCTGPTQSDCLACKN 223
QY 258 FNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLSLTDVGSCTLVCP 317
DB 224 FYDDGVCKQEGCPMGIQVNTYTFNPEPNDGKYAYGATCVRKCP-EHLIKONGACVRCPK 282
QY 318 HNQEVTAEADGTQRCCKSKPCAR 340
DB 283 GMPQNSE-----CVPGKGVCPK 300

Query Match 20.3%; Score 463.5; DB 13; Length 1137;
Best Local Similarity 44.8%; Pred. No. 1.3e-33;
Matches 86; Conservative 24; Mismatches 73; Indels 9; Gaps 5;

QY 161 LCYQDTILWKDIFHKNQLALTLDITNRSACHPCSPMCXGSRGWSSSDQSLRTVC 220
DB 3 LCPADTIHQDIIVRNFWASNFTLVPTNGSGGCRCHKCTG-RCWGPTEHNCQLTKTVC 61
QY 221 AGCG-ARCKGPLPTDCCHQCAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTF 279
DB 62 AEQDCGRGCVFYVSDCHRECGAGCGPKDTCFACFNFDNSGACVTCQGPQTFVNTTF 121
QY 280 ESMNPENEGRYTFGASCVTACPNYLSLTDVGSCTLVCPLEHNQEVTAEADGTQRCCKSKPC 339
DB 122 QLEHNHNAKYTYGAFCKGKCPHNFV-VDSSSCVRACFPSSKQEV-EENGINKWCKPCTDICP 179
QY 340 R-----GTHSL 346
DB 180 KACTDGTGTGSLV 191

Query Match 20.3%; Score 463.5; DB 13; Length 1137;
Best Local Similarity 44.8%; Pred. No. 1.3e-33;
Matches 86; Conservative 24; Mismatches 73; Indels 9; Gaps 5;

QY 161 LCYQDTILWKDIFHKNQLALTLDITNRSACHPCSPMCXGSRGWSSSDQSLRTVC 220
DB 3 LCPADTIHQDIIVRNFWASNFTLVPTNGSGGCRCHKCTG-RCWGPTEHNCQLTKTVC 61
QY 221 AGCG-ARCKGPLPTDCCHQCAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTF 279
DB 62 AEQDCGRGCVFYVSDCHRECGAGCGPKDTCFACFNFDNSGACVTCQGPQTFVNTTF 121
QY 280 ESMNPENEGRYTFGASCVTACPNYLSLTDVGSCTLVCPLEHNQEVTAEADGTQRCCKSKPC 339
DB 122 QLEHNHNAKYTYGAFCKGKCPHNFV-VDSSSCVRACFPSSKQEV-EENGINKWCKPCTDICP 179
QY 340 R-----GTHSL 346
DB 180 KACTDGTGTGSLV 191
```

RESULT 23

Q9W6F6

ID Q9W6F6 PRELIMINARY: PRT; 1137 AA.

RESULT 24
 Q9BG64 PRELIMINARY; PRT; 150 AA.
 AC Q9BG64; 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Receptor tyrosine kinase ErbB4 (Fragment).
 OS Oryctolagus cuniculus (Rabbit)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Kuechenhoff A.,
 RA Tetens P., Fischer B.,
 RT "ErbB genes and epidermal growth factor- (EGF-) like ligands in the
 RT peri-implantation rabbit uterus and blastocyst.",
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF333180; AAK14373.1;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0004714; F:transmembrane receptor protein tyrosine kin. . . ; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . . ; IEA.
 DR InterPro; IPR006211; Furin-like.
 DR InterPro; IPR006212; Furin repeat.
 DR Pfam; PF00757; Furin-like; 1.
 DR SMART; SM00261; FU; 2.
 FT NON TER 1
 FT NON TER 150
 SQ SEQUENCE 150 AA; 16491 MW; 56BFD01C70C610F5 CRC64;
 Query Match 17.7%; Score 404.5; DB 6; Length 150;
 Best Local Similarity 49.3%; Pred. No. 2.6e-29;
 Matches 73; Conservative 16; Mismatches 56; Indels 3; Gaps 3;
 QY 165 DTILMKDIFKHNQLALTIDNRSRACHPCSPWKGSRGWSSSDQSLSTVTCAGGC 224
 Db 2 DITHNQDIVRNPPSNLTIVSTNGSSGGRCHSKCTG-RWGPTEHCQTLTIVTCAGGC 60
 QY 225 -ARCKGRLPTDCHEQACAGCTGPKSDCLACLFHNSGICELHCPALVTYNTDTPESMP 283
 Db 61 DGRCYGPIVSDCHRECAAGCGSKPDTCFACMNFDSGACVTQCPQTFVYNTPTTQLEH 120
 QY 284 NPGRYTFGASCVTACPNYVLTVDGSC 311
 Db 121 NFNAKTYGAFVCKCPHFV-VDSGSC 147
 RESULT 25
 Q23821 PRELIMINARY; PRT; 1368 AA.
 AC Q23821;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DE Receptor tyrosine kinase.
 DE Receptor tyrosine kinase.
 GN LET-23 HOMOLOGUE.
 OS Caenorhabditis vulgaria.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pseudocercariae; Caenorhabditis.
 OX NCBI_TaxID=31233;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=9617760; PubMed=8604137;
 RA Sakai T., Koga M., Ohshima Y.;
 RT "Genomic structure and 5' regulatory regions of the let-23 gene in the
 RT nematode C. elegans."

J. Mol. Biol. 256:548-555(1996).
 RL EMBL; D63427; BAA09730.1; -.
 DR PIR; S70713; S70713.
 DR HSSP; P11362; 1FGK.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
 DR GO; GO:0004623; F:phospholipase A2 activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0016042; P:lipid catabolism; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . . ; IEA.
 DR InterPro; IPR000494; EGFR_L domain.
 DR InterPro; IPR006211; Furin-like.
 DR InterPro; IPR006212; Furin repeat.
 DR InterPro; IPR009030; Grow fac recep.
 DR InterPro; IPR01211; PhospholipaseA2.
 DR InterPro; IPR006162; Peptidase S.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR InterPro; IPR008266; Tyr_kinase_AS.
 DR Pfam; PF00757; Furin-like; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF01030; Recep_L domain; 2.
 DR PRINTS; PR00109; TYRKINASE.
 DR PRODOM; PR000001; Prot_kinase; 1.
 DR SMART; SM00261; FU; 6.
 DR SMART; SM00219; TyrcK; 1.
 DR PROSITE; PS00118; PA2_HIS; 1.
 DR PROSITE; PS00102; PHOSPHOPANTHETHEINE; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
 SQ SEQUENCE 1368 AA; 156391 MW; 65B087B8F6ACD85F CRC64;
 Query Match 17.6%; Score 402.5; DB 5; Length 1368;
 Best Local Similarity 28.1%; Pred. No. 6.8e-28;
 Matches 103; Conservative 59; Mismatches 136; Indels 69; Gaps 13;
 QY 25 VCTGDMKRLPASPETHLDMLHLYQCVQVQGNLELTPLTN----- 68
 Db 83 VCSGTNNLLSRYSNGNILEDL--EMYRCRRVYGNLEITWIEANIEIQWRNSTNOTVDAD 140
 QY 69 ----ASLGFLODIQEVQGVLIHQNVRQVPLQRLIRVGTQQLFEDNYALAVLDNGDPLN 124
 Db 141 IDYLKTVNFFDHLIELRSLIIRANIQKISPKLRVIYGVDEVFHDN-SLYIHQEK--- 196
 QY 125 NTTPVTGASPGGLREQLRSLTEILKGVLIQRPOLCQVDT-ILWKDIFKHNQLALT 183
 Db 197 -----VNELVKKELRVIRNGSVSIQNNPRMCFLATKVDWNEILYDXSRQKV- 243
 QY 184 IDNRSRACHPCSPM-----CKGSRGWSSSDQSLTRTVCCAGCARCKGLPT--- 233
 Db 244 -XNKHKACWNGELIASXHENCK-DKCGKGDNDCKIYRSVCPCPCQCFYSNITQSY 301
 QY 234 DCCHQCAAGCTGPKRHSCLACLFHNSGICELHCPALVTYNTDTPESNPEGRYTFGA 293
 Db 302 ECDSSCLGGCTHGPDSACIACKYEMDEMCIIDTCFARKIFNHKTGRVLPNDGRYQNG 361
 QY 294 SVCTACPNYLTVDGSCVLCPLHNSQVTAEDGT---QRCEKC-SKPCARGTHSLPRP 349
 Db 362 HCVCCKCPPELLIXND-----VCRHCSGHHYDATDMRECEKCPSSGC-----PKI 408
 QY 350 AAVPVPL 356
 Db 409 CTVDGFL 415
 RESULT 26
 Q9BG65 PRELIMINARY; PRT; 151 AA.
 ID Q9BG65

Q26565; 2001 (TrEMBLrel. 17, Created)
 AC 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Receptor tyrosine Kinase ErbB3 (Fragment).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Kuechenhoff A.,
 RA Tetens F., Fischer B.;
 RA "ErB genes and epidermal growth factor- (EGF-) like ligands in the
 RA peri-implantation rabbit uterus and blastocyst";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF333179; AAK14372.1; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0007169; P:protein amino acid phosphorylation; IEA.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; IEA.
 DR InterPro; IPR00494; EGFR_L domain.
 DR InterPro; IPR006211; Furin-like.
 DR InterPro; IPR009030; Grow_fac_recep.
 DR Pfam; PF00757; Furin-like; 1.
 DR Pfam; PF01030; Recep_L domain; 1.
 DR SMART; SM00261; FU; 1.
 DR Receptor; Signal. 19 POTENTIAL.
 KW SIGNAL 20 366 POTENTIAL.
 FT CHAIN 20 366 POTENTIAL.
 SQ SEQUENCE 366 AA; 41153 MW; 5052971783019A8 CRC64;
 Query Match 15.9%; Score 363.5; DB 5; Length 366;
 Best Local Similarity 29.9%; Pred. No. 5e-25;
 Matches 100; Conservative 46; Mismatches 129; Indels 59; Gaps 13;
 QY 36 PASPEHLMRLHLYGCGQVGNLELTLP-----TNASLSFLQDIQVGVYLIHNO 90
 DB 42 ENPKHILQTYIKFLYGGCTHIIENLVICGLEKLENGSDPDLSPLEKIDVSGYVYIGNS 101
 QY 91 VROVPLQRLIVRGTFQTFEDNYALVLDNGDPLNTPVTGASPGGLRELQLSLTEILK 150
 DB 102 VKTISLPSLVKVG-----EPGYRIMTSAALVISRNSLEILDLRLSLTAIOR 148
 QY 151 GGVLIOIRNFC-LYQDTILWOLFHNQO-----LATLIDTVRSR- 190
 DB 149 NDVALNNOFLCNFGFTIDWEQIFEDNRKQMTIPDRKEKTVSHAGCDIALRKVTDRTKH 208
 QY 191 ACHPCSPMKCG-SRCWGESSEDCQSLTR-----TVGAGGCARCKGFLPTDCHEQCAAG 243
 DB 209 SCHGSCPVVNGRGYCWGPRKEMCKMLKCANPNKYCLGSA-----TTPCLEELGG 262
 QY 244 C-TGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNPEGYTTCASCVTHCPYN 302
 DB 263 CETRP--GNCRACKHAMNDGKCVSQCPPLIVSRRESRTVANPEFKYNFHDICVKNCPAP 320
 QY 303 YLSTDVSGSTLCPLHNOEVTAEADGTQCEKCSK 336
 DB 321 FLKSD-SYCVIECDLNTQ--IPVNGI--CKDCPK 349
 RESULT 28
 Q26566 PRELIMINARY; PRT; 1717 AA.
 AC Q26566;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Epidermal growth factor receptor precursor.
 GN SER.
 OS Schistosoma mansoni (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
 OC Schistosomatidae; Schistosomatidae; Schistosoma.
 NCBI_TaxID=6183;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Puerto Rican;
 RC MEDLINE=92365727; PubMed=1501637;
 RA Shoemaker C.B., Ramachandran H., Landa A., dos Reis M.G., Stein L.;
 RT "Alternative splicing of the Schistosoma mansoni gene encoding a
 RT homologue of epidermal growth factor receptor.";
 RL Mol. Biochem. Parasitol. 53:17-32(1992).
 DR EMBL; M86396; AAA29866.1; -
 DR PIR; A45558; A45558.
 DR HSP; P11362; IFGK.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.

Q26565; 2001 (TrEMBLrel. 17, Created)
 AC 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Receptor tyrosine Kinase ErbB3 (Fragment).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Kuechenhoff A.,
 RA Tetens F., Fischer B.;
 RA "ErB genes and epidermal growth factor- (EGF-) like ligands in the
 RA peri-implantation rabbit uterus and blastocyst";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF333179; AAK14372.1; -
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0007169; P:protein amino acid phosphorylation; IEA.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; IEA.
 DR InterPro; IPR00494; EGFR_L domain.
 DR InterPro; IPR006211; Furin-like.
 DR InterPro; IPR009030; Grow_fac_recep.
 DR Pfam; PF00757; Furin-like; 1.
 DR Pfam; PF01030; Recep_L domain; 1.
 DR SMART; SM00261; FU; 1.
 DR Receptor; Signal. 19 POTENTIAL.
 KW SIGNAL 20 366 POTENTIAL.
 FT CHAIN 20 366 POTENTIAL.
 SQ SEQUENCE 366 AA; 41153 MW; 5052971783019A8 CRC64;
 Query Match 17.3%; Score 395; DB 6; Length 151;
 Best Local Similarity 45.4%; Pred. No. 2e-28;
 Matches 74; Conservative 28; Mismatches 47; Indels 14; Gaps 5;
 QY 94 VLOQLRIVRGTFQTFEDNYALVLDNGDPLNTPVTGASPGGLRELQLSLTEILKGV 153
 DB 1 LPLFNLAIVRGTYDGKFAIFVM-----LNYNT---NSSHALRQLRQLTEILSGV 51
 QY 154 LIQRNPOLCYDITLWIDIFHNQALTLTIDTVRSACHPCSPKCGSRWGESSEDDQ 213
 DB 52 YIEKNDKLCMDTIDWDIV---RDPGAEIVVKNRGRSCPPCHEVCCKG-RWGSFGPDDQ 107
 QY 214 SLRTVCAGGC-ARCKGFLPTDCHEQCAAGCTGPKHSDCLAC 255
 DB 108 TLITTCAPQNGHCFGDPDNCQCHDEACGCGSGPDTDCFAC 150
 RESULT 27
 Q26569 PRELIMINARY; PRT; 366 AA.
 AC Q26569;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Epidermal growth factor receptor precursor.
 GN SER.
 OS Schistosoma mansoni (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
 OC Schistosomatidae; Schistosomatidae; Schistosoma.
 NCBI_TaxID=6183;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Puerto Rican;
 RC MEDLINE=92365727; PubMed=1501637;
 RA Shoemaker C.B., Ramachandran H., Landa A., dos Reis M.G., Stein L.;
 RT "Alternative splicing of the Schistosoma mansoni gene encoding a
 RT homologue of epidermal growth factor receptor.";
 RL Mol. Biochem. Parasitol. 53:17-32(1992).
 DR EMBL; M86399; AAA29869.1; -
 DR PIR; D45558; D45558.

```
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006466; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR InterPro; IPR000494; EGFR_L domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow fac recep.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008466; Tyr_kinase.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00757; Furin-like; 2.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF01030; Recep_L domain; 2.
DR PRINTS; PR00109; TYRKINASE
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 8.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR ATP-binding; Kinase; Receptor; Signal; Transferase;
KW Tyrosine-protein kinase.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1717 EPIDERMAL GROWTH FACTOR RECEPTOR.
SQ SEQUENCE 1717 AA; 192304 MW; 1101A338C1653D8C CRC64;

Query Match 15.9%; Score 363.5; DB 5; Length 1717;
Best Local Similarity 29.9%; Pred No. 3.6e-24;
Matches 100; Conservative 46; Mismatches 129; Indels 59; Gaps 13;

QY 36 PASEPETHLMLRLHYQGVQVQGNLELYLP-----TNASLFLQDIOEVGYVLIH 90
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 42 PNPXHIQLTYIKFLYGGCTHIIGNLVICGLEKLENGSDPLSFLEKIEDVSGVYVIGNS 101
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 91 VRQVPLQRLIRVGRQPLFEDNYALVLDNGPLNNTTPTVTCASPGGLRELQRLSLTELK 150
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 102 VKTSLPSLKVIRG-----EPGYRIMNTSAALVIRNSULEILERSLTAIOR 148
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 151 GGVLIQRLPQLC-YQDTILWKIDIFHNKQ-----LALFLIDNRSR- 190
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 149 NDIVALLNQFLNCFDTIDWQIFEDNRKQMFIDRKETVSHAGCQIALRKYYTDDRTHK 208
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 191 ACHPCSPMCKG-SRWGSESEDCSLTR-----TVAGGCARCKGPIPTCCHQCAAG 243
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 209 SCHGSCFVYNGRGYCWGPKPMCKQKLCANPNPDNYCLGGR-----TTQPCLEECIG 262
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 244 C-TGPKSDCLACLFHNHSGICELHCPALVTYNTDTESMPNPSERYTEGASCVTACPN 302
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 263 CETRP--GNCRCAXEMNDGKVCQCPPLIVSRESERTVANPEKYNFHDICVNCAP 320
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 303 YLSTDVSGCTLVCPHMQVEAEDGTQRCCKSK 336
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 321 FLKSD-SYCVIECDLNTQ--IPVNGT--CKDCPK 349

RESULT 29
QY1X8
ID QY1X8 PRELIMINARY; PRT; 1193 AA.
AC QY1X8;
DT 01-NOV-1999 (TRENBlrel. 12, Created)
DT 01-MAR-2002 (TRENBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBlrel. 25, Last annotation update)
DE Protein tyrosine kinase.
GN EPTK178.
OS Ephydria fluviatilis.
OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;
OC Haplosclerida; Spongillidae; Ephydria.
OX NCBI_TaxID=31330;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=92464375; PubMed=10229568;
RA Suga H., Koyanagi M., Hoshiyama D., Ono K., Iwabe N., Kuma K.,
RA Miyata T.;
RT "Extensive gene duplication in the early evolution of animals before
RT the paracon-eumetazoan split demonstrated by G proteins and protein
RT tyrosine kinases from sponge and hydra.";
RL J. Mol. Evol. 48:646-653(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21601119; PubMed=11738833;
RA Suga H., Katoh K., Miyata T.;
RT "Sponge homologs of vertebrate protein tyrosine kinases and frequent
RT domain shufflings in the early evolution of animals before the
RT paracon-eumetazoan split.";
RL Gene 280:195-201(2001)
DR EMBL; AB006570; BAA81724.2; --
DR HSPF; P08631; IAD5.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005234; F:ATP binding; IEA.
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006466; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR InterPro; IPR000494; EGFR_L domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow fac recep.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008466; Tyr_kinase.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 6.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00652; TNFR_NGFR_1; 1.
KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 1193 AA; 128169 MW; 009E4AC9BC12DF60 CRC64;

Query Match 15.9%; Score 353.5; DB 5; Length 1193;
Best Local Similarity 30.1%; Pred. No. 1.9e-23;
Matches 113; Conservative 42; Mismatches 155; Indels 65; Gaps 19;

QY 4 AALCRWGLLI-----ALLPRG--AASQVCTGTDMLRLPASPETH-LDMLRLHYQGVQ 55
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 3 SFFCLFSLLEAPCCIALSAPNCVIGTACRGNAANTFVQGSSTSEVLQALARFCGCTR 62
QY 56 VQGNLELYLP-----TNASLFLQDIOEVGYVLIH-NQVRQVPLQRLIRVRG-TQ 105
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 63 ISGNLAI-HLNQLEDANLTEDSFASLVHLREVSQYLFQSPNIPAVTRITLPNLAVIRQTA 121
QY 107 LPEDNYALVLDNGDPLNNTTPTVTCASPGGLRELQRLSLTELKGGVLIQRLPQLCQD 166
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 122 LSSNTNGVSVLLVENSIRGS-----VNLPSLSTEISNGGATFVSTNGMCGFLG 167
QY 167 ILWKDIFHNKQ-----ALFLIDNRSRACHPCSPMCKGSRCKGSESEDCQSLRTV 219
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 168 VNWADIL-TNGQLDYSMSGKVTLD-----CS-SCPSGHGW-SSPHYCQTLTKTV 215
QY 220 CAGGCARC---KQPLPTDCHEOCACGCTGPKHSDCLACLFHNHSGICELHCPALVTYNT 276
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 216 CPQCGCRCFQSPGDSGYQCCSALCAAGCSGPTSDQCYACSLDNNGTCTVTECPAQVYDP 275
QY 277 DTPESMPNPSERYTFGASCVTACPNVLSLDVSGCTLVCPHMQVEAEDGTQRCCKSK 336
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 276 DLFWLVDPNPNRYLASDGLCVQCCAGNLLLEY-AGACVTFCP---EGYTSNNGA-KVCACSS 330
```

Job time : 79.8815 secs

```
Qy 337 -PCAR-----GTHSL 345
Db 331 GVCPRVCTGTGTGDL 345

RESULT 30
Q26567 PRELIMINARY; PRT; 334 AA.
AC Q26567;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Epidermal growth factor receptor precursor.
GN SER.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatoidea; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Puerto Rican;
RX MEDLINE=92365727; PubMed=1501637;
RA Shoemaker C.B., Ramachandran H., Landa A., dos Reis M.G., Stein L.;
RT "Alternative splicing of the Schistosoma mansoni gene encoding a
RT homologue of epidermal growth factor receptor.";
RL Mol. Biochem. Parasitol. 53:17-32(1992).
DR EMBL; M86397; AAA29867.1; -.
DR PIR; B45558; B45558.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:Arp binding; IEA.
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; IEA.
DR InterPro; IPR000494; EGFR_L domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow_fac recep.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L domain; 1.
DR SMART; SMO0261; FU; 1.
KW Receptor; Signal.
FT SIGNAL 19 POTENTIAL.
FT CHAIN 20 334 POTENTIAL.
SQ SEQUENCE 334 AA; 37747 MW; 55848013A48292E CRC64;

Query Match 14.5%; Score 331; DB 5; Length 334;
Best Local Similarity 29.9%; Pred. No. 4.4e-22;
Matches 89, Conservative 40; Mismatches 115; Indels 54; Gaps 10;

Qy 36 PASPEHLDMLRHLYQSCQVQGNLELYLP-----TNASLSFLQDIQFVQGVYLAHQ 90
Db 42 PNPKHILQTYIKFLYGGCTHIGNLVICGLEKLENGSDPDLFLKIEDVSGYVYIGONS 101
Qy 91 VRQVPLQRLTRVGTQFEDNYALVLDNGDPLNNTPTVTGASPGGLRELQLRSLTEILK 150
Db 102 VKTISLPSLKVIRG-----EPGYRINWNTSAALVISRNSLEIIDLRLSLTAIOR 148
Qy 151 GGVLIQRNQPC-YQDTILMKDIFHKNQ-----LALTLDITNSR- 190
Db 149 NDIVALNNQFLCNFGFTIDWEQIFEDNRKQMFIPDRKEKTVSHAGCDIALKRYTDDRKH 208
Qy 191 ACHPCSPCKG-SRCWGESSEDCQLTR-----TVGAGGCARCKGLPTDCHEQCAAG 243
Db 209 SCHGSCPVVNRGVCWGPCKPMQCMKLCANNPNYCLGGR-----TTQPCLECLGG 262
Qy 244 C-TGPKHSDCLACLFHNSGICELHCPALVTYNTDTTFESMPNPEGRYTFGASCVTACP 300
Db 263 CETRP--GNCRACKHAMNDKVCQCPPLIVSREGRVANPEKYNFHDICVKNCP 318
```

Search completed: July 4, 2004, 04:21:00

This Page Blank (uspto)

This Page Blank (uspto)